Small Angle Neutron Scattering (SANS) Studies on R67 Dihydrofolate Reductase, a Tetrameric Protein with Intrinsically Disordered N-Termini

Purva P. Bhojane[†], Michael R. Duff, Jr. [†], Khushboo Bafna[‡], Pratul Agarwal[§], Christopher Stanley[‖] and Elizabeth E. Howell[†,‡,*]

[†]Department of Biochemistry, Cellular and Molecular Biology, University of Tennessee, Knoxville, TN, 37996-0840

[‡]Genome Science and Technology Program, University of Tennessee, Knoxville, TN, 37996-0840

[§]Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee, 37831

[‖]Biology and Soft Matter Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee, 37831

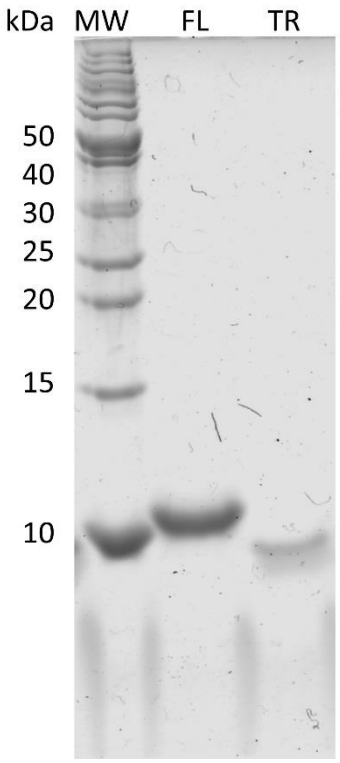


Figure S1. 20% SDS-PAGE gel of the truncated R67 DHFR after 24 hours of digestion by chymotrypsin-agarose at room temperature. The gel is labelled to indicate the wells that contain the molecular weight marker (MW), the full length R67 DHFR (FL), and the truncated protein (TR). Molecular weights of the standard ladder are also indicated to the left of the gel.

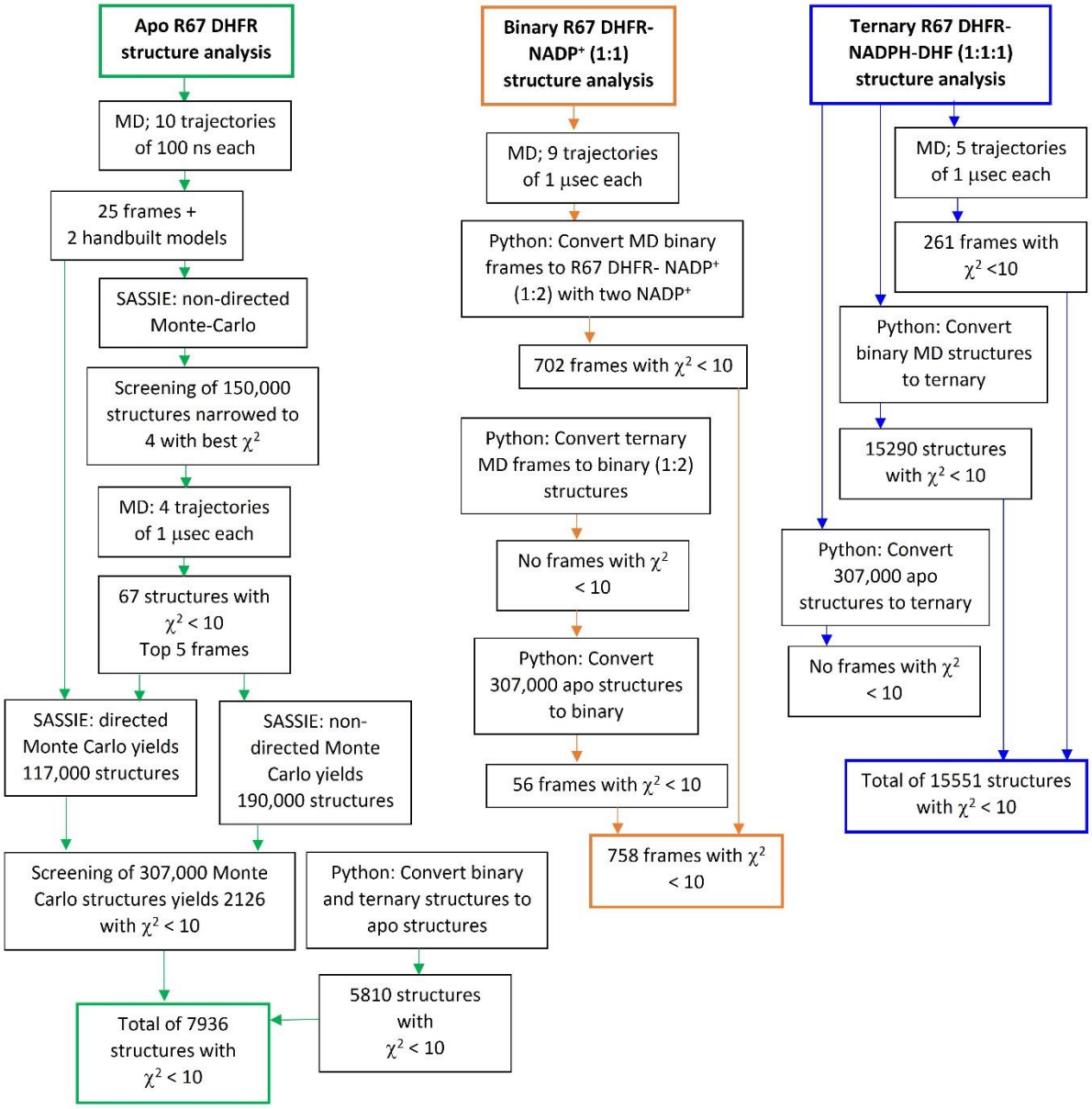


Figure S2. Flow chart for SASSIE analysis. The workflow for the apo, binary and ternary data analyses in SASSIE are shown indicating the source of the conformers that fit the respective SANS data. Many different steps were necessary to achieve a sufficient number of good fits.

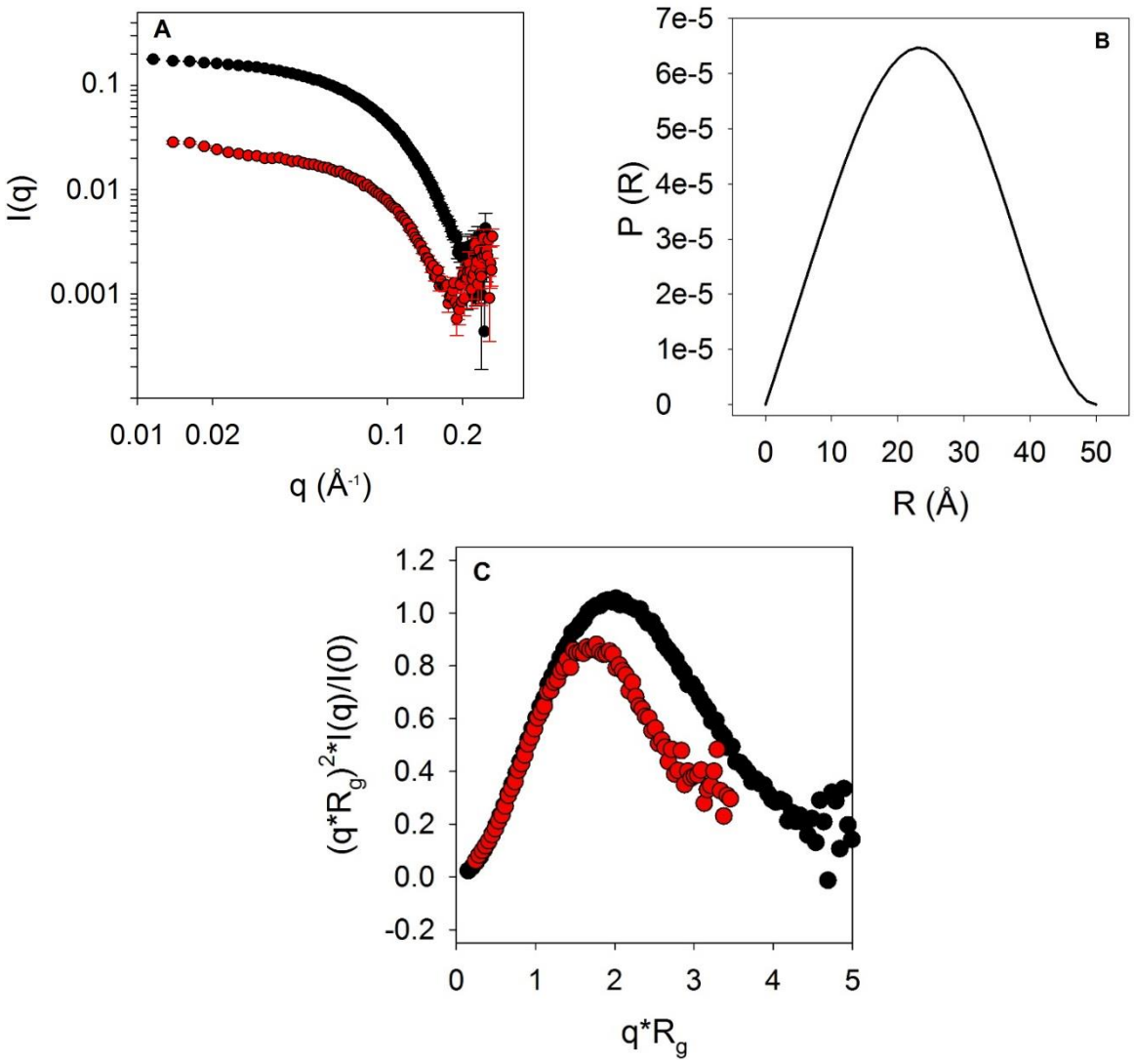


Figure S3. SANS profile and GNOM analysis for truncated R67 DHFR. Panel A shows the SANS scattering profile for truncated R67 DHFR (●). Data for full length R67 DHFR (●) are shown for comparison. A GNOM fit of the profile for truncated R67 DHFR, shown in panel B, gives the pairwise distance distribution and an R_g value of 17.86 ± 0.14 Å. Dimensionless Kratky plots of the full length and truncated proteins indicate that both are globular as shown in panel C. The collapse of the N-termini on the core of R67 in the full length protein may make it

hard to differentiate between the compact truncated protein and the full length R67 with disordered N-termini.

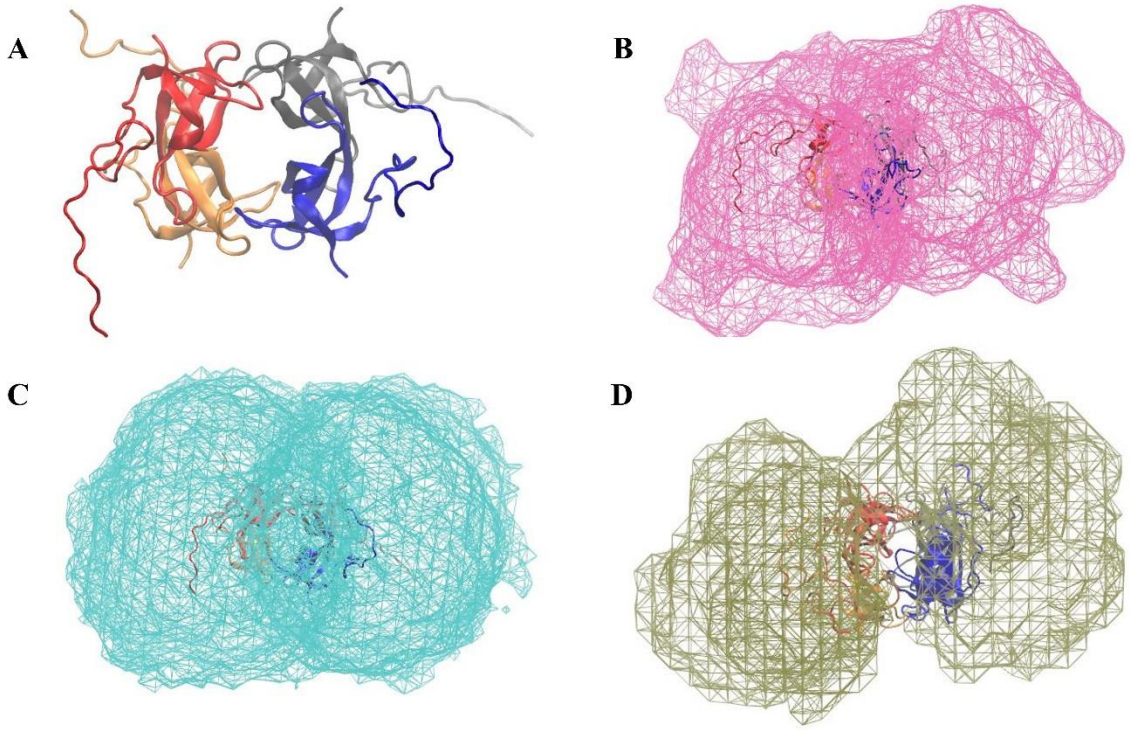


Figure S4. Structural space of the MD and Monte Carlo data sets used in our studies. A) The Modeller structure used as the initial starting structure in the MD and Monte Carlo studies. Each different monomer is shown in a different color. B) The volumetric mesh obtained from molecular dynamics simulations of the apo, binary and ternary complexes. C) Mesh plots from both complex and directed Monte Carlo analysis of apo R67 DHFR. D) The conformational space obtained from our two additional models where the N-termini were either adjusted to interact with each other on both sides of the pore, or were moved to cover access to the active site pore.

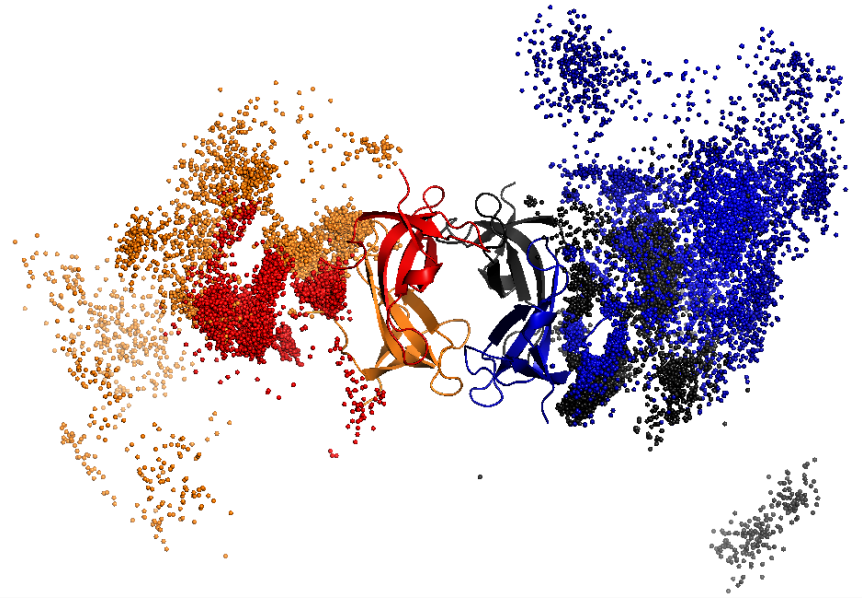
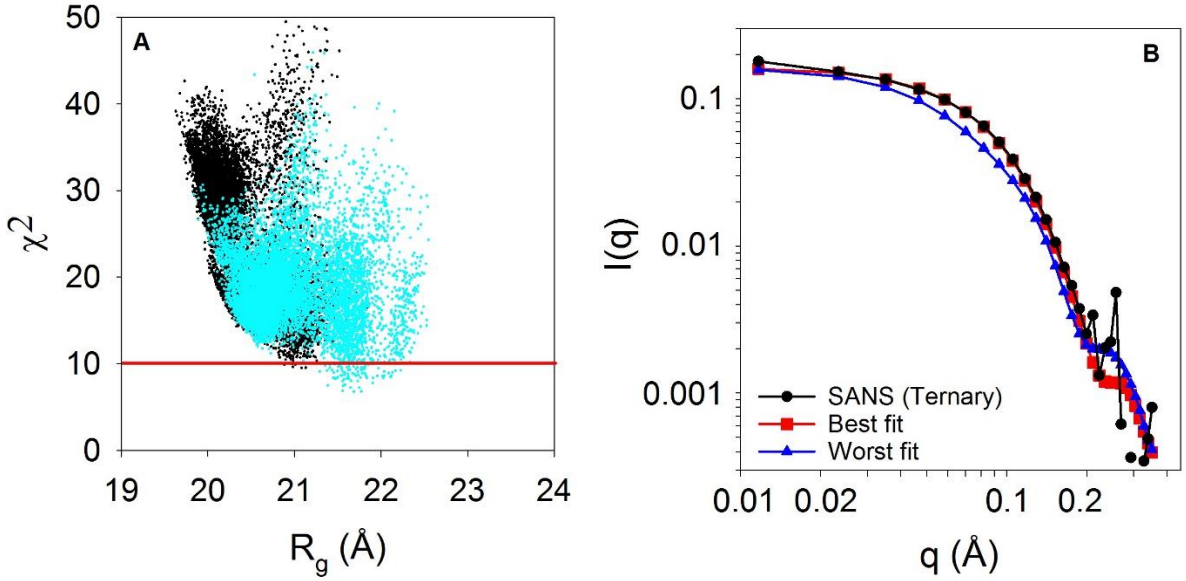
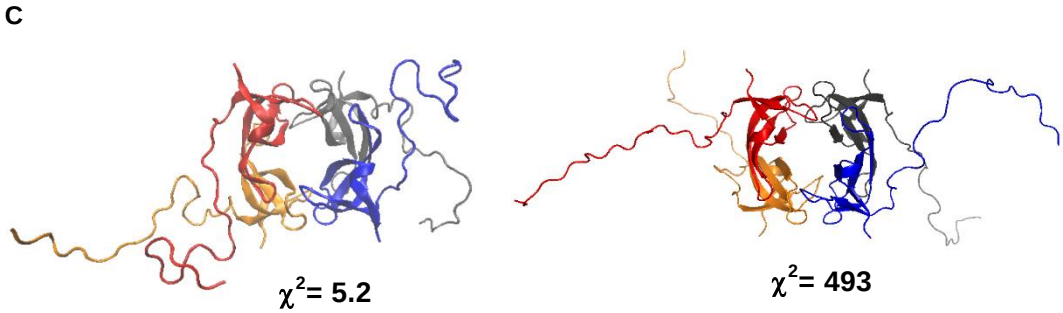
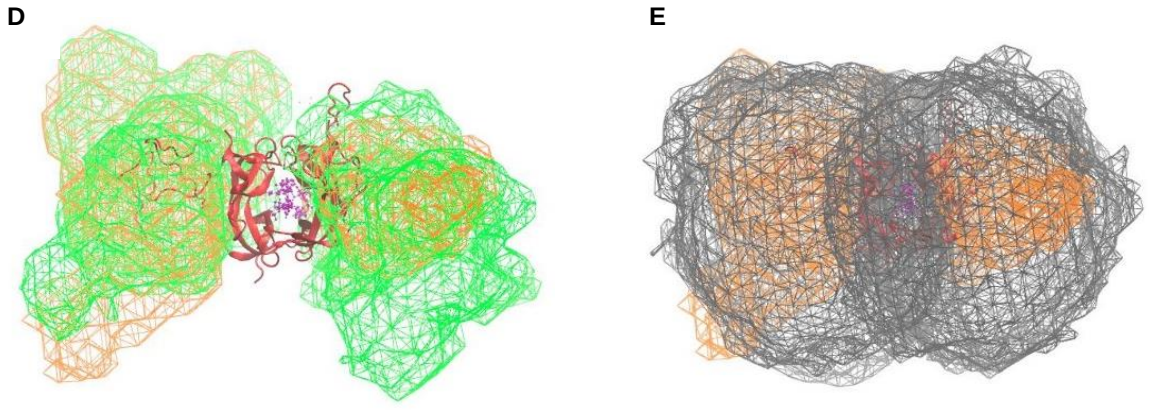


Figure S5. Analysis of apo, full length R67 DHFR showing the center of mass points for the N-terminal methionine residues for each of the four N-termini of the good fits from SASSIE with a χ^2 <10. The four monomers are colored differently and the center of mass points are color coded for each monomer.


A
50
40
30
20
10
0
χ2
19
20
21
22
23
24
Rg (Å)
B
0.1
0.01
0.001
I(q)
SANS (Ternary)
Best fit
Worst fit
0.01
0.02
0.1
0.2
q (Å)


C
χ2= 5.2
χ2= 493


D
E

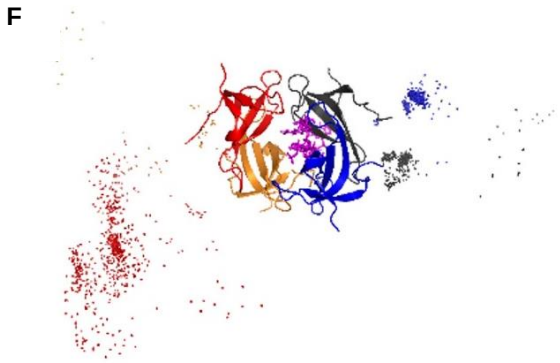


Figure S6. SASSIE analysis of the disordered N-termini upon $NADP^+$ binding to full length R67 DHFR. Panel A shows the resulting χ^2 vs. R_g plot generated by MD (•) and Monte-Carlo (•) sampling. The red line indicates a cut off for χ^2=10. Panel B overlays the theoretical and experimental SANS profiles for the best and worst fits (lowest and highest χ^2 values respectively). Panel C shows the corresponding structures for the best and worst fits. In panels D and E, the orange mesh depicting the density plot for all the good fits (758) to the binary complex data is overlaid on the density plot of all the frames that fit to the apo data (green mesh in panel D and all the sampled structures (dark gray mesh in panel E). Bound $NADP^+$ (magenta) molecules are shown as ball-and-stick models. Panel F represents the center of mass points of the N-terminal methionine residues for each of the four N-termini of the fits from SASSIE with a χ^2 <10. The four monomers are colored differently and the center of mass points are color coded for each monomer.

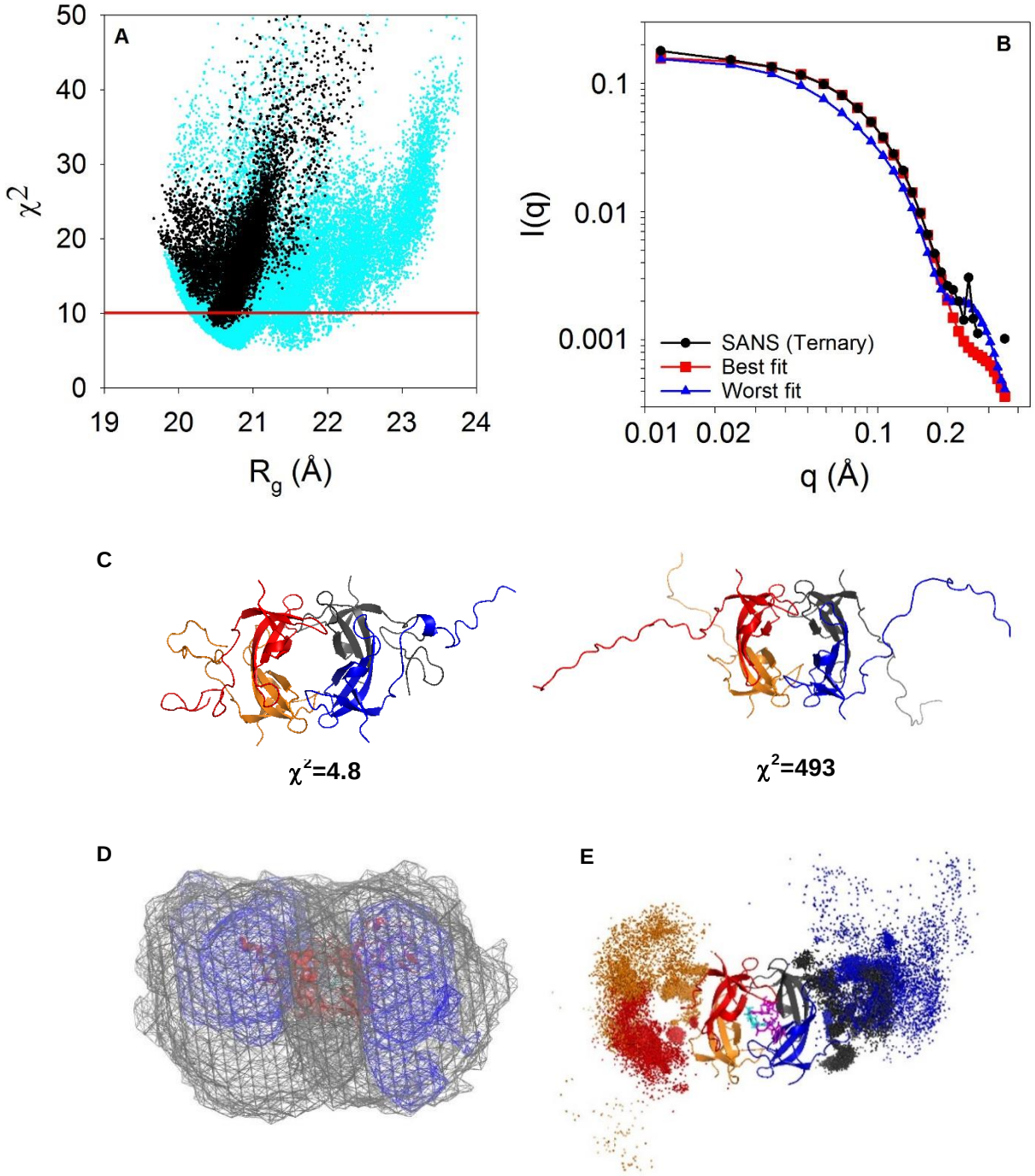
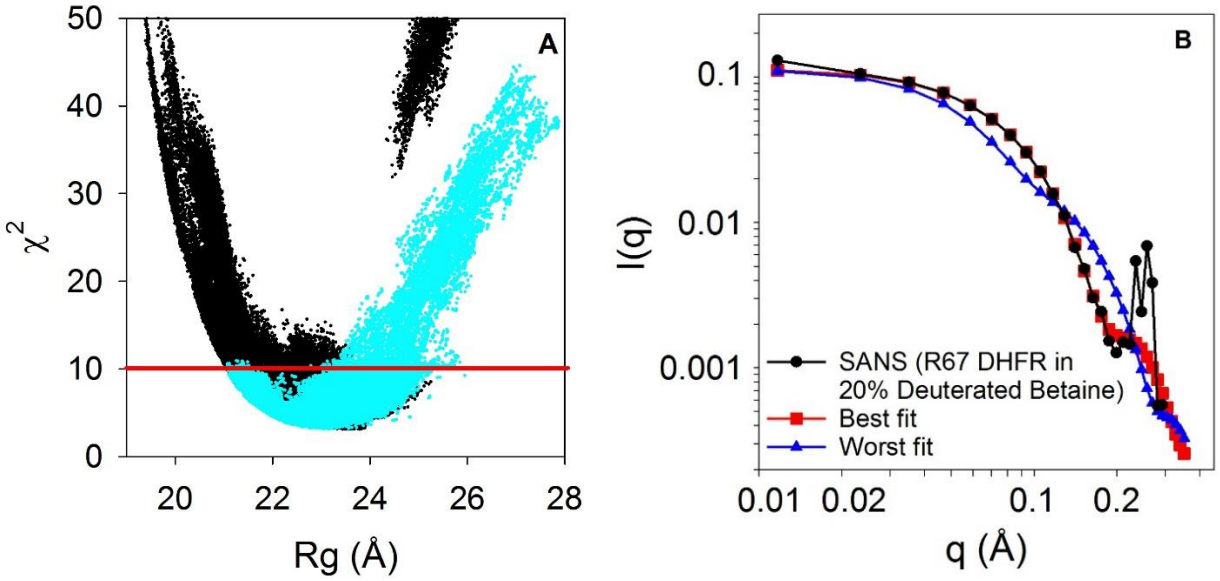


Figure S7. Conformational analysis of the disordered N-termini upon DHF binding to the R67 DHFR-$NADP^+$ complex by SASSIE. The χ^2 vs. R_g plot shown in panel A compares the

representative frames generated by MD (•) and Monte-Carlo (•) sampling to the experimental SANS data for R67 DHFR-$NADP^+$-DHF ternary complex. The red line indicates a cut off of χ^2 =10. An overlay of the theoretical and experimental SANS profiles for the best (χ^2 = 4.8) and worst (χ^2 = 493) fits, respectively) is shown in panel B. The corresponding best and worst frames are shown in panels C and panel D, representing the overlay of the density plot for the 15,551 frames that fit the ternary data (blue mesh) on the entire structural space sampled (dark gray mesh). The ligands DHF (cyan) and NADPH (magenta) are shown in the active site pore of panel D. Panel E represents the center of mass points of the N-terminal methionine residue for each of the four N-termini of the best frames from SASSIE with a χ^2 <10. The four monomers are colored differently and the center of mass points are color coded for each monomer.

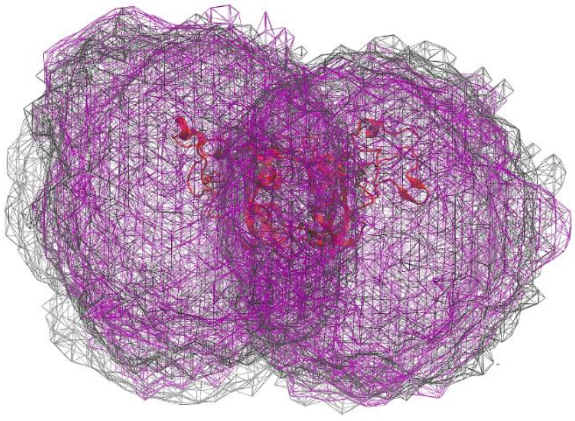


C

χ^2=3.14; R_g = 23.79 Å

χ^2=140.9; R_g = 29.58 Å

D



E

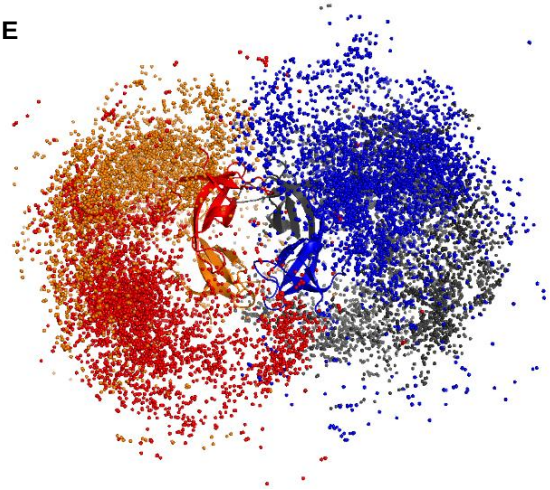


Figure S8. Conformational analysis of the disordered N-termini of R67 DHFR in the presence of 20 % deuterated betaine. Panel A shows the χ^2 vs. R_g plot that compares the experimental SANS data for apo R67 DHFR in the presence of 20 % deuterated betaine with the theoretical SANS profiles from structures obtain with MD (•) and Monte-Carlo (•). The red line indicates a cut off for χ^2=10. Panel B shows an overlay of the theoretical and experimental SANS profiles for the best and worst fits. The corresponding structures are shown in panel C. Panel D shows the overlay of the density plot representing 58,277 frames that fit to the SANS for apo R67 DHFR in the presence of 20 % deuterated betaine (purple mesh) on the entire structural space sampled (dark gray mesh). Panel E shows the center of mass of the first residue of each of the N-termini associated with good fits to the experimental data. The 4 chains and center of mass points are color coded. Out of 58,277 frames, the center of mass is shown for 15,000 representative frames.

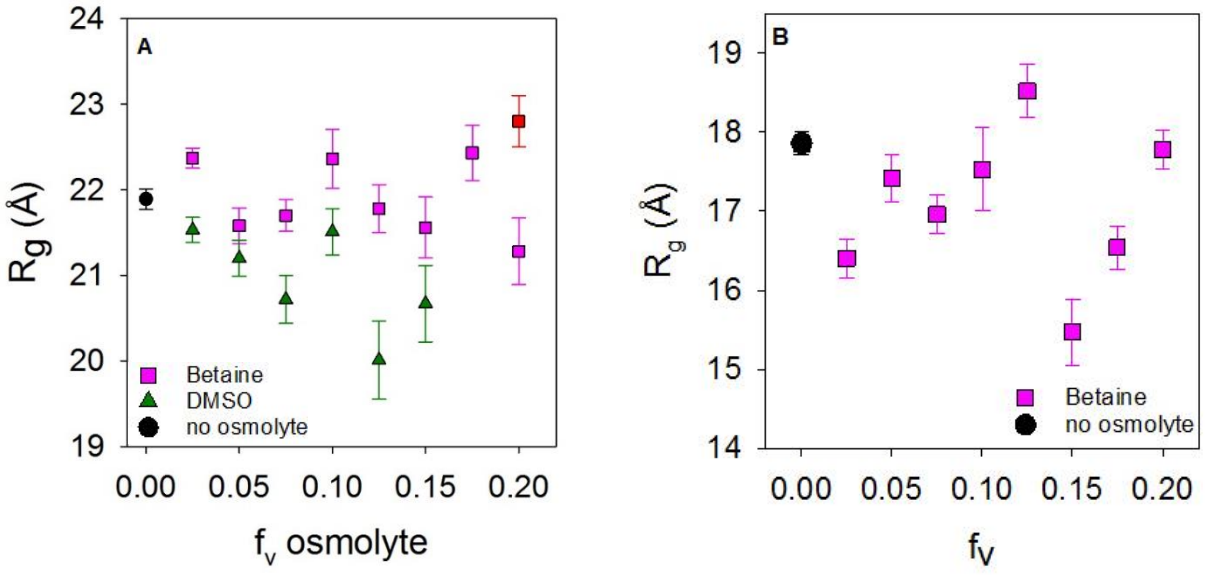
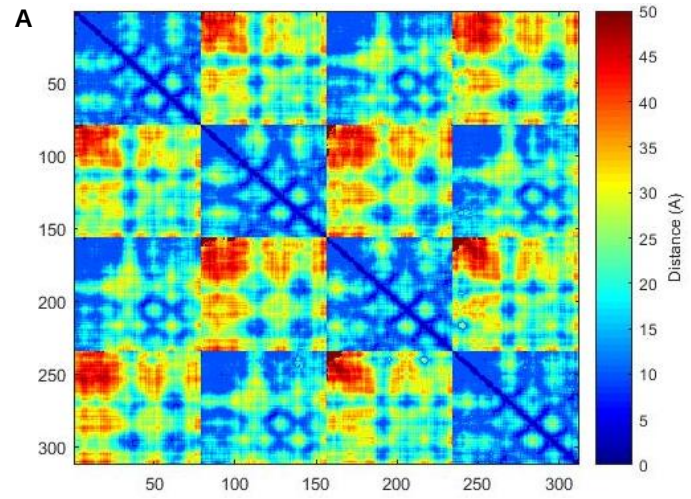
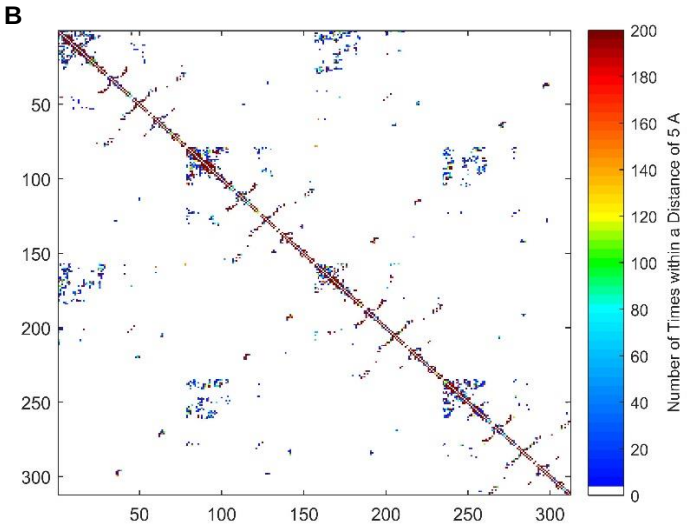


Figure S9. SANS profiles of R67 DHFR in the presence and absence of osmolytes and the analysis of the radii of gyration. Panel A shows the overlay of the SANS profiles at ~ 6 mg/mL R67 DHFR in 20 mM deuterated Tris buffer in D_2O (pD 7.0) with no osmolyte (●) and with increasing fractional volume, f_v, of betaine (■) and DMSO (▲). The radii of gyration of 21.89 ± 0.12 Å for no osmolyte, 21.28 ± 0.38 Å with 20 % betaine and 20.67 ± 0.45 Å with 15 % DMSO were obtained from the GNOM analysis. Additionally, panel A shows the average R_g value obtained from the SASSIE analysis that yield 58,277 protein conformers that fit to the SANS data for R67 DHFR in the presence of 20 % deuterated betaine (■). Panel B shows the variation in R_g for R67 DHFR with increasing fractional volume, f_v, of betaine (■) and DMSO (▲).


A
50
100
150
200
250
300
Distance (A)
0
5
10
15
20
25
30
35
40
45
50


B
50
100
150
200
250
300
Number of Times within a Distance of 5 A
0
20
40
60
80
100
120
140
160
180
200

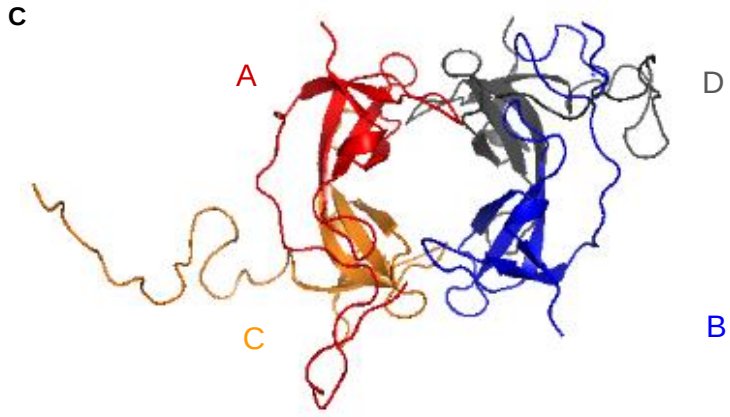

C
A
D
C
B

Figure S10. Data mining of the apo protein SASSIE conformers. Panel A shows the heat map for the minimum distance between the centers of mass for the residue pairs. The x and y axes are the residues in all four chains numbered 1-312. The symmetry of the structure can be seen by the repeating patterns. Panel B (same x and y axes) shows the frequency of the closely positioned centers of mass of residues within a distance of 5 Å. Most interactions are seen within residues of the same chain (along the diagonal) as well as the N-terminal residues on the same side of the core of the protein, i.e. between chains A-C and chains B-D (i.e. off-diagonal points). Panel C shows a representative model for the full-length R67 DHFR with the four monomers labelled A-D.

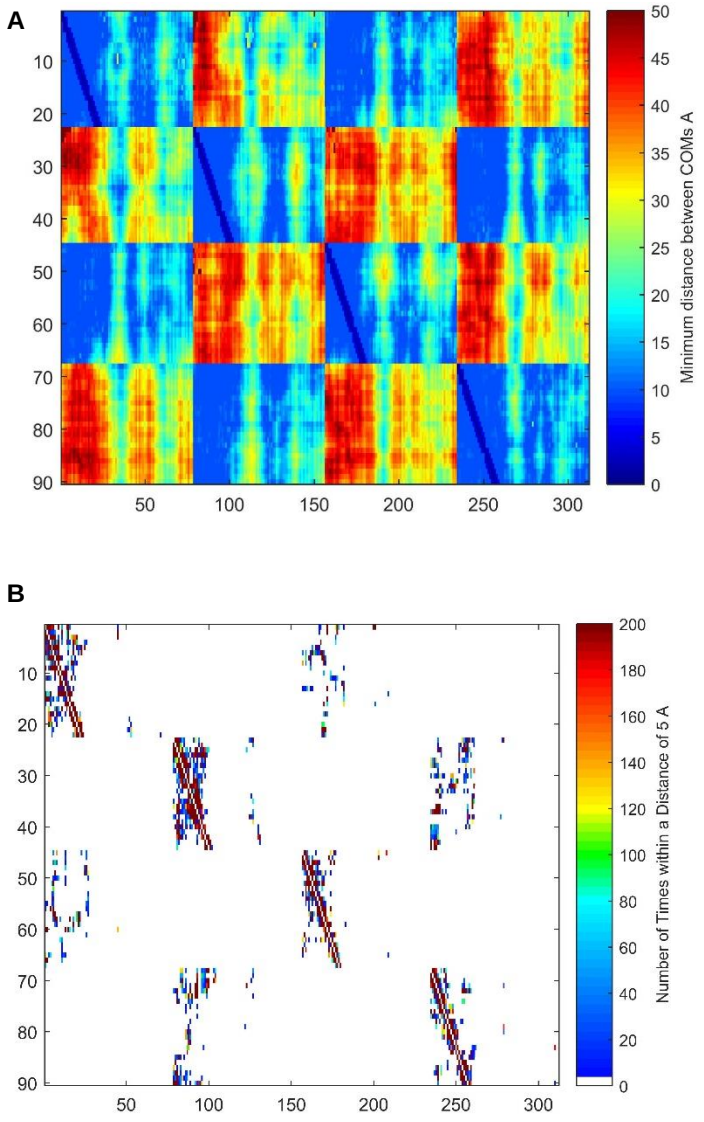


Figure S11. Data mining analysis for the ternary complex. Panel A shows the heat map for the minimum distance between the centers of mass for the residue pairs. Residues in all four chains are numbered 1-312 on the x-axis while the y-axis shows only the N-terminal 21 residues for the four chains. Amino acids 1-21 describe the N-terminal residues in monomer A, residues 22-43 describe the same 21 N-terminal residues in monomer B, residues 44-65 describe the same 21 N-terminal residues of monomer C, while residues 66-87 describe the same 21 N-terminal residues

of monomer D. Panel B shows the frequency of the closely positioned centers of mass of residues within a distance of 5 Å. Many interactions are seen near the diagonal that describe intra-chain interactions. Inter-chain interactions occur off the diagonal on the same side of the protein pore, i.e. between chains A-C and chains B-D. See panel C of Figure S10 for a representative model for the full-length R67 DHFR with the four monomers labelled A-D.

Table S1 – Comparison of R_g values obtained from the data analysis for the full length and truncated R67 DHFR using Guinier analysis and the GNOM program.

Full length R67 DHFR									
Betaine	**0 %**	**2.5 %**	**5 %**	**7.5 %**	**10 %**	**12.5 %**	**15 %**	**17.5 %**	**20 %**
GNOM R_g (Å)	21.89 ± 0.12	22.37 ± 0.12	21.58 ± 0.21	21.7 ± 0.19	22.36 ± 0.34	21.78 ± 0.28	21.56 ± 0.36	22.43 ± 0.32	21.28 ± 0.39
Guinier R_g (Å)	21.62	21.94	21.99	21.54	22.33	21.42	22.07	22.45	21.01
DMSO	**0 %**	**2.5 %**	**5 %**	**7.5 %**	**10 %**	**12.5 %**	**15 %**	**17.5 %**	**20 %**
GNOM R_g (Å)	21.89 ± 0.12	21.53 ± 0.15	21.2 ± 0.21	20.72 ± 0.28	21.51 ± 0.27	20.01 ± 0.46	20.67 ± 0.45	-	-
Guinier R_g (Å)	21.62	20.67	21.19	19.51	21.57	19.42	19.43	-	-
Truncated R67 DHFR									
Betaine	**0 %**	**2.5 %**	**5 %**	**7.5 %**	**10 %**	**12.5 %**	**15 %**	**17.5 %**	**20 %**
GNOM R_g (Å)	17.86 ± 0.14	16.40 ± 0.25	17.42 ± 0.3	16.96 ± 0.24	17. 53 ± 0.52	18.51 ± 0.34	15.47 ± 0.42	16. 54 ± 0.27	17.78 ± 0.25
Guinier R_g (Å)	17.85	17.51	17.76	17.36	18.05	13.56	13.67	14.88	15.16

Table S2 - Comparison of melting temperatures of full length and truncated R67 DHFR with and without osmolytes. The thermal denaturation studied by DSC fits to a three-state model that yields two melting temperatures.

Protein	Osmolyte	TM_1 (° C)	TM_2 (° C)
Full Length R67 DHFR	None	66.8 ± 0.2	68.7 ± 0.1
	20% Betaine	68.9 ± 0.1	71.7 ± 0.1
	15% DMSO	58.2 ± 0.2	61.3 ± 0.1
Truncated R67 DHFR	None	59.9 ± 0.2	63.8 ± 0.1
	20% Betaine	64.0 ± 0.3	67.7 ± 0.3
	15% DMSO	52.8 ± 0.3	56.7 ± 0.7