Supporting Information for: Interpretation of solution X-ray scattering by explicit-solvent molecular dynamics

Po-chia Chen[1] and Jochen S. Hub[1*]

[1] Institute for Microbiology and Genetics
Georg-August-University Göttingen
Justus-von-Liebig-Weg 11
37077 Göttingen, Germany

[*] E-mail: jhub@gwdg.de; phone: +49 551 39-14189

Simulation Details

Choice and preparation of source coordinates

Apo and holo structure of LBP were taken from protein data bank (PDB codes 1USG and 1USI, respectively) (1).

The CTP-bound ATCase from [PDB code 1ZA1 (2)] and ATP-bound ATCase from [PDB code 7AT1 (3)] are commonly accepted to represent the crystallographic T and R states, respectively. Bound ATP and CTP were retained during all simulations, but the substrate mimic PALA was removed from the R coordinates. Although it is necessary *in-vivo* to form the R state, we observed no collapse back to T during 50 ns-simulations (see main text).

CRM1 simulations were based on the ring-open structure from *chaetomium thermophilum* CRM1 [ctCRM1, PDB code 4FGV (4)], as an appropriate model for mammalian CRM1 in the absence of binding partners (40% sequence identity). ctCRM1 represents the most complete structure and also retains a C-terminal helix stretching across the ring. In comparison, the structure of mouse CRM1 (PDB code 3GJX) (5) was not suitable due to non-trivial rearrangement upon ranGTP binding, while the structure of human CRM1 (PDB code 3GB8) (6) lacks several long loops and the N-terminal HEAT repeat and would require speculative modeling. Missing loops near the N-terminus of ctCRM1 were modelled using the kinematic loop modeling module from the Rosetta3.4 suite (7).

Simulation setup and parameters

Equilibrium simulations were conducted using the GROMACS 4.6 simulation software (8). SWAXS-driven MD simulations were conducted using our in-house modification of GROMACS 4.6. Hydrogen atoms were added with the pdb2gmx software (8), keeping co-crystallised water molecules and the ligands noted above. The proteins were solvated in a rhombic-dodecahedron box with 15$\sim$10 Å buffer regions, and ionised to 100 mM NaCl, if not stated otherwise. CRM1 simulations without salt were neutralized by Na^+ counterions. Parameters for LBP and ATCase simulations were taken from the CHARMM27

and CHARMM22* force fields, respectively (9–11). CRM1 was simulated either with the Amber99sb or with the CHARMM22* force field, as noted in the main text (12). Water was described by the TIP3P water model (13).

Electrostatic interactions were computed using the particle-mesh Ewald algorithm (14). Short-range dispersive interactions were modeled by a Lennard-Jones potential using dual-range cut-off between 10 and 12 Å using force-switching. Water molecules were kept rigid using the SETTLE algorithm (15), and all other bonds were constrained using P-LINCS (16). Hydrogen bonds were described by virtual interaction sites, allowing a time step of 4 fs (17). The pressure was kept at 1 bar using the Berendsen barostat during initial equilibrations and the Parrinello-Rahman barostat during equilibrium simulations (time constant 1 ps) (18, 19). Likewise, the temperature was controlled at 298.15 K using velocity-rescaling during equilibration and using Nosé-Hoover thermostat during equilibrium simulations (time constant 5 ps)(20, 21). Protein and non-protein atoms have been coupled to separate thermostats. Stochastic dynamics integrators (22) were used during SWAXS-driven MD simulations because SWAXS-driven MD does not conserve energy, which might otherwise results in gradual heating of the system.

Free (unbiased) simulations

Each system was equilibrated over 2 ns using stepwise decreasing restraints to relax the solute, first releasing the sidechains and then the backbones. From each equilibrated system, a number of free (*i.e.*, not SWAXS-driven) simulation were conducted, as listed in Table S1. SWAXS curves from equilibrium simulations were computed as explained previously (23). The principal component analysis (PCA) for LPB was conducted on the C_α atoms, after combining free 100-ns simulations of apo and phenylalanine-bound LBP.

SWAXS-driven simulations

Target experimental patterns for ATCases have been transcribed from Fetler *et al.* (24), while the target data for ctCRM1 were provided by the Ralf Ficner group. In the latter case, sample and buffer intensities were obtained to conduct and verify background matching in-house. To reduce the influence of noise upon WAXSMD simulations, the buffer-subtracted curve was passed through a 20-point running average. This smoothed curve was used in SWAXS-driven MD. For modeling the solution ensemble of CRM1 as a combination of open and closed states, the unsmoothed curve was adopted, allowing for direct comparison between our *chi*-metric and χ^2_{free} validations according to Rambo and Tainer (25). In this work, we adopt convention $I_{\text{sample}} - I_{\text{buffer}}$ for buffer subtraction.

Before coupling the simulations to an experimental SWAXS curve, the experimental curves were fitted to a SWAXS curve calculated from a short equilibrium simulation. Accordingly, the experimental curve was fitted by minimizing

$$\chi^2 = \sum_i \left[\log I_c(q_i) - (f \, \log I_e(q_i) + c) \right]^2, \tag{S1}$$

where f adjusts the arbitrary overall scale of scattering curves, and c is an offset that approximately absorbs experimental uncertainties due to the buffer subtraction. Fitting only these two parameters, we

recently found excellent agreement between calculated and experimental WAXS curves (23). Here, the fit was conducted on a log scale, compatible to the coupling potential E_{SWAXS} in eq. 2.

Parameters applied for SWAXS-driven MD are listed in Table S1. A second simulation for ATCase $R \rightarrow R$ was conducted at a higher k_c to resolve the discrepancy at $2\,\mathrm{nm}^{-1}$, producing the blue ensemble in the main text and Fig. S9. As noted the main text, the simulation frames were superimposed onto a reference structure in reference orientation (ω) before evaluating the averages in eqs. 7, 8, and 13 (in main text), using a least-square fit on the C_α atoms. For each protein, the same reference structure was used for all SWAXS-driven simulations.

The spatial envelope for LBP was constructed around simulation frames after combining 100-ns trajectories of apo and phenylalanine-bound LBP, keeping a distance of 0.7 nm between protein and envelope. The envelope was constructed as explained previously (23). Because the maximum extent of CRM1 and ATCase were not known prior to SWAXS-driven simulation, we assigned spherical envelopes with radii 5.2 and 8.8 nm, respectively, based on free simulations and solvation layer thickness of 0.7 nm. A snapshot of each system with its envelope are shown in Figure S1. The same envelope was shared across all SWAXS calculations of the same system.

Rigid-body modeling

Rigid body modeling was conducted with SASREF(26) starting either from apo- or holo- LBP crystal structures, and targeting the computed SAXS curve of either open or closed LBP ensembles in Fig. 1. This resulted in four sets of models, colored as plus symbols in Fig. 2). The N-and C-terminal domains of LBP were defined as follows: the N-terminal domain spanning D1-L120 and P249-K327, and the C-terminal domain spanning D121-M248 and G332-K346. Four linker residues were removed between K327 and G332. To constrain the possible quaternary arrangements, constraints at 3, 4, 5, 6 and 7 Å (in steps of 1 Å) were imposed between L120-D121 and P249-M248. Five replicates were conducted for each constraint value, resulting in a total of 25 modeling attempts per structure-SAXS combination.

P(r) calculations with GNOM

The program GNOM (27) was used to back-calculate the pair distribution function P(r) from simulated and experimental curves. For each protein and curve, a range of maximum extent R_{max} values is chosen to compute candidate P(r) curves. The curve with the highest fitness score calculated by GNOM was chosen, subject to the condition that it did not exhibit a large region of $\sim$0 density. These resultant R_{max} values are also reported in Table S2. In order to maintain comparability with experimental data, the maximum-q of the simulated SWAXS curves considered for P(r) calculations for ATCase and ctCRM1 were set a 5 nm^{-1}. The cut-off for LBP was set at 10 nm^{-1}.

Computational resource usage

All simulations were conducted on the GWDG cluster at the Georg-August University Göttingen, and at the North-German Supercomputing Alliance (HLRN) cluster in Hanover using between 16 to 192 CPU cores depending on architecture and system size. SWAXS-driven MD simulations take approximately

10% longer than conventional MD simulations. According to GROMACS internal cycle-accounting, the SWAXS-related calculations took (relative to total computing time) $7.6 \pm 1.4\%$ during LBP, $11.7 \pm 2.2\%$ during ATCase, and $9.9 \pm 2.0\%$ during ctCRM1 simulations.

Supporting References

1. Magnusson, U., B. Salopek-Sondi, L. A. Luck, and S. L. Mowbray, 2004. X-ray Structures of the Leucine-binding Protein Illustrate Conformational Changes and the Basis of Ligand Specificity. *J. Biol. Chem.* 279:8747–8752.

2. Wang, J., K. A. Stieglitz, J. P. Cardia, and E. R. Kantrowitz, 2005. Structural basis for ordered substrate binding and cooperativity in aspartate transcarbamoylase. *PNAS* 102:8881–8886.

3. Gouaux, J. E., R. C. Stevens, and W. N. Lipscomb, 1990. Crystal structures of aspartate carbamoyltransferase ligated with phosphonoacetamide, malonate, and CTP or ATP at 2.8-.ANG. resolution and neutral pH. *Biochemistry* 29:7702–7715.

4. Monecke, T., D. Haselbach, B. Voß, A. Russek, P. Neumann, E. Thomson, E. Hurt, U. Zachariae, H. Stark, H. Grubmüller, A. Dickmanns, and R. Ficner, 2013. Structural basis for cooperativity of CRM1 export complex formation. *Proc. Natl. Acad. Sci. USA* 110:960–965.

5. Monecke, T., T. Güttler, P. Neumann, A. Dickmanns, D. Görlich, and R. Ficner, 2009. Crystal Structure of the Nuclear Export Receptor CRM1 in Complex with Snurportin1 and RanGTP. *Science* 324:1087–1091.

6. Dong, X., A. Biswas, K. E. Süel, L. K. Jackson, R. Martinez, H. Gu, and Y. M. Chook, 2009. Structural basis for leucine-rich nuclear export signal recognition by CRM1. *Nature* 458:1136–1141.

7. Mandell, D. J., E. A. Coutsias, and T. Kortemme, 2009. Sub-angstrom accuracy in protein loop reconstruction by robotics-inspired conformational sampling. *Nat. Methods* 6:551–552.

8. Hess, B., C. Kutzner, D. van der Spoel, and E. Lindahl, 2008. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* 4:435–447.

9. Brooks, B. R., C. L. Brooks, A. D. Mackerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, and M. Karplus, 2009. CHARMM: The biomolecular simulation program. *J. Comput. Chem.* 30:1545–1614.

10. Piana, S., K. Lindorff-Larsen, and D. E. Shaw, 2011. How Robust Are Protein Folding Simulations with Respect to Force Field Parameterization? *Biophys. J.* 100:L47–L49.

11. Bjelkmar, P., P. Larsson, M. A. Cuendet, B. Hess, and E. Lindahl, 2010. Implementation of the CHARMM Force Field in GROMACS: Analysis of Protein Stability Effects from Correction Maps, Virtual Interaction Sites, and Water Models. *J. Chem. Theory Comput.* 6:459–466.

12. Hornak, V., R. Abel, A. Okur, B. Strockbine, A. Roitberg, and C. Simmerling, 2006. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* 65:712–725.

13. Jorgensen, W. L., J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–935.

14. Essmann, U., L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen, 1995. A smooth particle mesh Ewald method. *J. Chem. Phys.* 103:8577–8593.

15. Miyamoto, S., and P. A. Kollman, 1992. SETTLE: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comput. Chem.* 13:952–962.

16. Hess, B., 2008. P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation. *J. Chem. Theory Comput.* 4:116–122.

17. Feenstra, K. A., B. Hess, and H. J. C. Berendsen, 1999. Improving efficiency of large time-scale molecular dynamics simulations of hydrogen-rich systems. *J. Comput. Chem.* 20:786–798.

18. Parrinello, M., and A. Rahman, 1981. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* 52:7182–7190.

19. Berendsen, H. J. C., J. P. M. Postma, W. F. v. Gunsteren, A. DiNola, and J. R. Haak, 1984. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* 81:3684–3690.

20. Bussi, G., D. Donadio, and M. Parrinello, 2007. Canonical sampling through velocity rescaling. *J. Chem. Phys.* 126:014101.

21. Hoover, W. G., 1985. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A* 31:1695–1697.

22. Van Gunsteren, W. F., and H. J. C. Berendsen, 1988. A Leap-frog Algorithm for Stochastic Dynamics. *Mol. Simul.* 1:173–185.

23. Chen, P.-c., and J. S. Hub, 2014. Validating Solution Ensembles from Molecular Dynamics Simulation by Wide-Angle X-ray Scattering Data. *Biophys. J.* 107:435–447.

24. Fetler, L., and P. Vachette, 2001. The Allosteric activator Mg-ATP Modifies the Quaternary Structure of the R-state of Escherichia coli Aspartate Transcarbamylase Without Altering the T↔R Equilibrium. *J. Mol. Biol.* 309:817–832.

25. Rambo, R. P., and J. A. Tainer, 2013. Accurate assessment of mass, models and resolution by small-angle scattering. *Nature* 496:477–481.

26. Petoukhov, M. V., and D. I. Svergun, 2005. Global Rigid Body Modeling of Macromolecular Complexes against Small-Angle Scattering Data. *Biophys. J.* 89:1237–1250.

27. Svergun, D. I., 1992. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.* 25:495–503.

Supplementary Tables

Table S1. Simulation parameters for free ensembles. Replicas are reported per functional state. Initial t_{eq} ns discarded as equilibration.

Variables	LBP	ATCase	bs-RNase	ctCRM1
replicas	10 (ensemble),1 (WAXS)	3	5	10
t_{sim} [ns]	50,100	T: 20, R: 40	100	50
t_{eq} [ns]	10	T: 0, R: 20	20	20

Table S2. Simulation Parameters for SWAXS-driven MD, and structural information used to assist in selection of scattering vectors. Maximum extent, R_{max}, taken from GNOM calculations, and converted to width of a Shannon channel $q_{Shannon} = \pi/R_{max}$, and total number of Shannon channels within coupling range $n_{Shannon}$. Scattering vectors coupled in SWAXS-driven MD are distributed linearly between q_{min} and q_{max}, resulting in a corresponding separation q_{sep} that is directly comparable to $q_{Shannon}$. *: Additional ATCase simulations with single-q-point restraints are conducted with increased k_c.

System	LBP	ATCase	ctCRM1
R_{max} [nm]	6.2	13	11
$q_{Shannon}$ [nm^{-1}]	0.51	0.24	0.29
$n_{Shannon}$	15.8	9.1	8.4
q_{sep} [nm^{-1}]	0.5	0.2	0.2
n_q	16	11	12
q_{min} [nm^{-1}]	0.5	0.2	0.4
q_{max} [nm^{-1}]	8.0	2.2	2.6
k_c	100∼3000	500 (2500)*	1000
τ [ns]	1.0	1.0	2.5
δt [ps]	2.0	2.0	2.0
Total simulation time [ns]	10	20∼23	40

Supplementary Figures

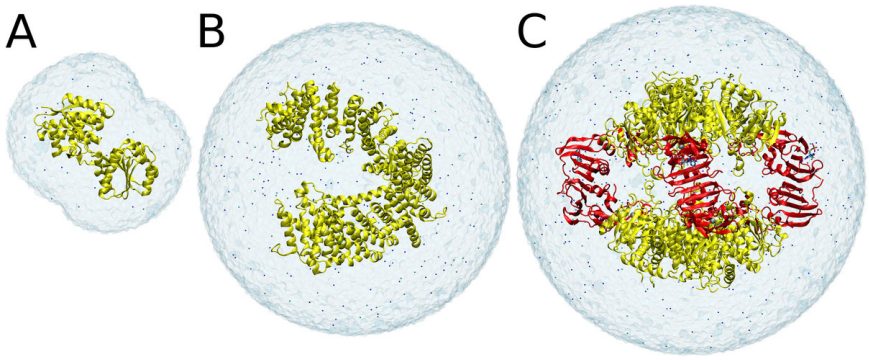

Figure S1. Biological systems presented in this paper, displayed to the same scale. The solute and bound ligands (if any) are shown in cartoon and stick formats, respectively. Solvating ions are shown as small spheres, while water molecules are abstracted for clarity. Each solute is enclosed in an envelope, in which all atoms are considered for SWAXS scattering calculations (see Methods of the main text). (A) LBP in the apo-state. (B) ctCRM1 in its ring-open state. (C) ATCase in the *R*-state, with bound ATP. Catalytic trimers and regulatory dimers are colored yelow and red, respectively.

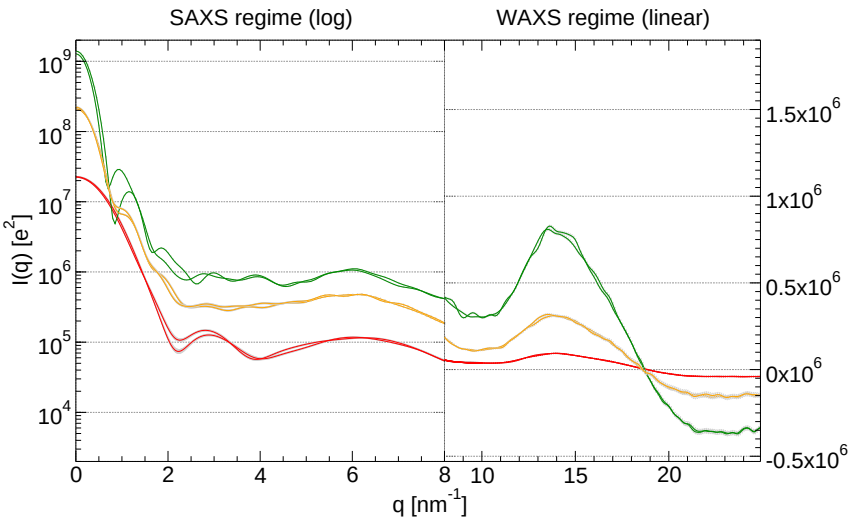

Figure S2. SAXS and WAXS patterns of the biological systems considered in this study, superimposed without rescaling to indicate relative intensities. Left panel shows $\log I(q)$ over $q < 8\,\mathrm{nm}^{-1}$, in the SAXS/near-WAXS regime, while the right panel shows $I(q)$ over $8\,\mathrm{nm}^{-1} < q < 25\,\mathrm{nm}^{-1}$ in the WAXS regime. Error bars (where visible,) are shown in grey. Colours are as follows: Red – LBP in the apo- (open) and holo- (closed) states. Orange – ctCRM1 simulated under CHARMM22* forcefield, in position restrained (ring-open) and free (ring-closed) simulations. Green – ATCase in the R- and T- states, in free simulations. The curves show that the intensities mainly change in the SAXS regime upon conformational transition, but hardly in the WAXS regime, justifying why we here coupled SWAXS-driven MD only to SAXS and near-WAXS data.

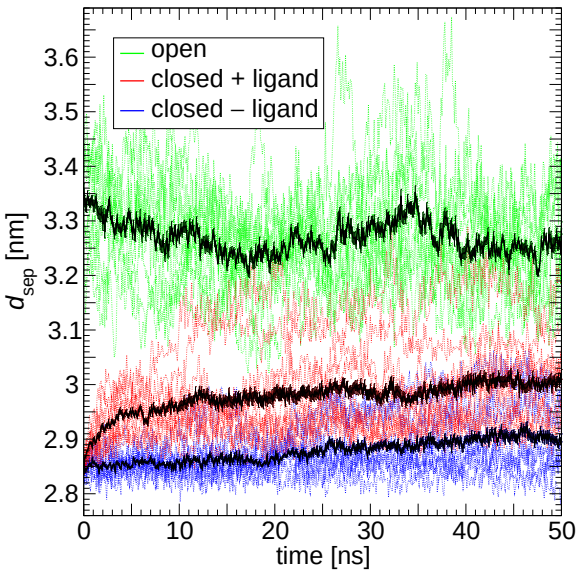

Figure S3. Domain separation d_{sep} during free 50ns-simulations of LBP, starting from open (green), closed (red), and closed with ligand removed (blue) configurations. Colored, dotted lines represent individual trajectories, and thick black lines represent average over ten replicates.

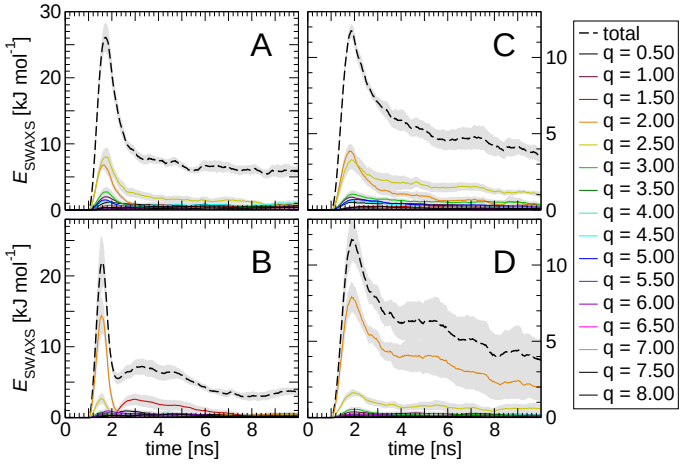

Figure S4. Decomposition of E_{SWAXS} during LBP SWAXS-driven MD simulations, demonstrating that the conformational transitions are mainly driven by variations of the intensity at $q = 2\,\mathrm{nm}^{-1}$. The total E_{SWAXS} (thick black, dotted) is plotted along with contributions from individual scattering angles (colored as shown in legends). All values represent averages over 5 replicate trajectories, with standard errors shown as a solid light-grey background. (A/C) open LBP targeting SWAXS curves of closed LBP, using (A) $k_c = 1000$ or (C) $k_c = 300$. (B/D) closed LBP targeting SWAXS curves of open LBP, using (B) $k_c = 1000$ or (D) $k_c = 300$.

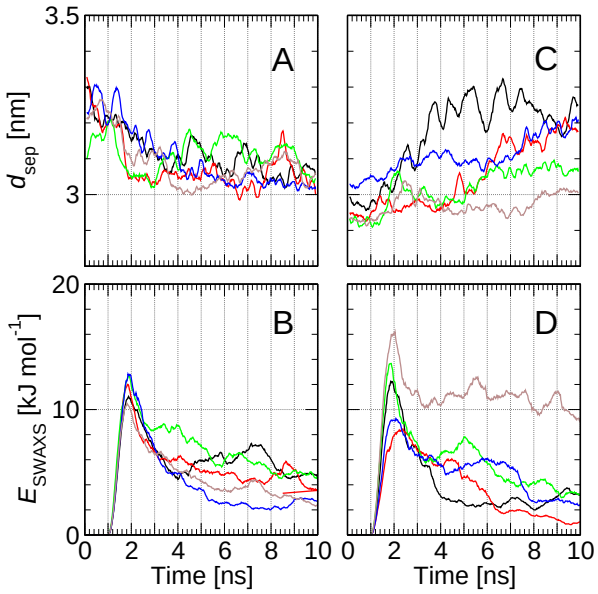

Figure S5. Opening and closing trajectories of SWAXS-driven MD simulations of LBP at $k_c = 300$ and $\tau = 1$ ns. (A/C) smoothed domain separation d_{sep} and (B/D) SWAXS potential E_{SWAXS} versus simulation time. (A/B) d_{sep} and E_{SWAXS} for closing transforms; (C/D) d_{sep} and E_{SWAXS} for opening transforms.

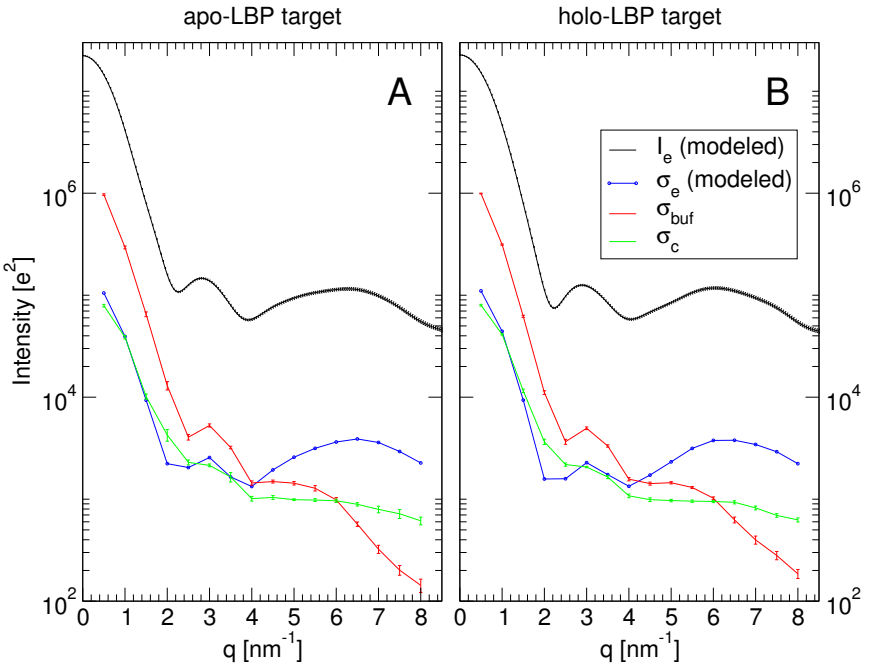

Figure S6. Statistical and systematic uncertainties applied as inverse weights for SWAXS-driven MD of LBP (eq. 3 of the main text). (A) Targeting the apo state, and (B) the holo state of LBP; black: Target SWAXS curves; blue: experimental error, modeled as a linear increase from 0.5% at $q = 0$ to 5% at $q = 8\,\mathrm{nm}^{-1}$ of I_e, resembling the statistical error in the experimental data used in a previous study(23); red: systematic error in $I_c(q)$ due to an uncertainty of 1% in the buffer density; green: statistical error of $I_c(q)$. The overall uncertainty is dominated by the buffer at low q, and by the experimental error at high q.

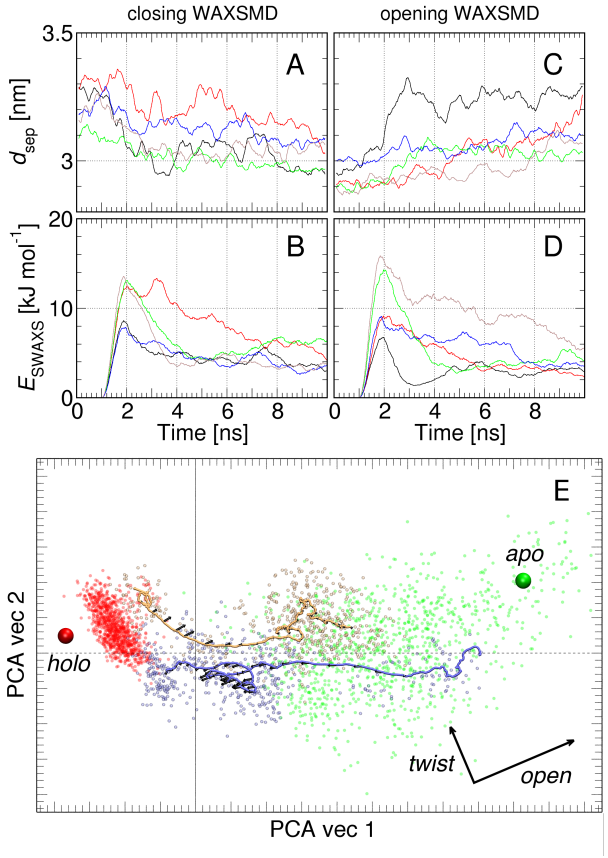

Figure S7. Opening and closing trajectories of SWAXS-driven MD simulations of LBP at $k_c = 1$ and $\tau = 1$ ns, using the errors as inverse weights instead of uniform weighting (eq. 3 of the main text, and Fig. S6). (A/C) smoothed domain separation d_{sep} and (B/D) SWAXS potential E_{SWAXS} versus simulation. (A/B) for closing transitions; and (C/D) for opening transitions. (E) PCA analysis of the black trajectory in (A-D), using the same color scheme as Fig. 2. The green and red spheres indicated the apo and holo crystal structures, respectively.

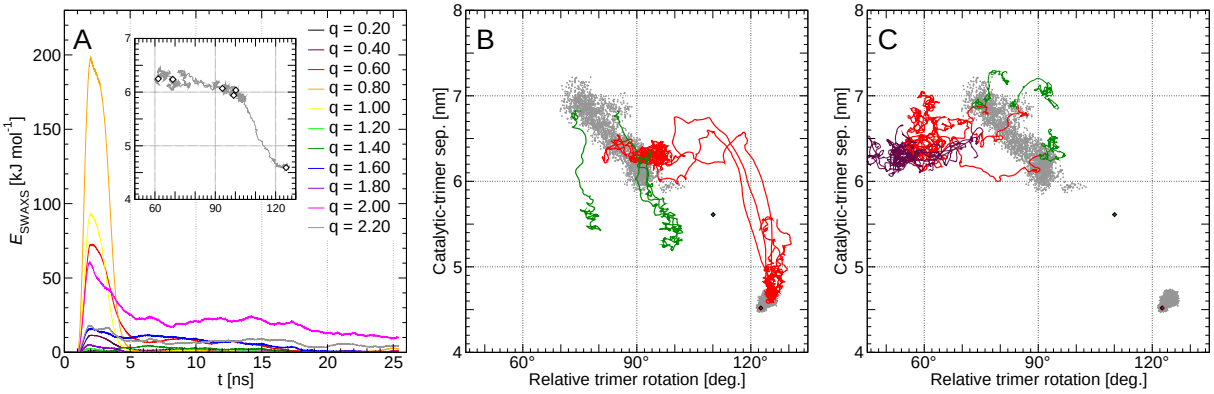

Figure S8. (A) Decomposition of E_{SWAXS} into contributions from different scattering vectors q, shown for one typical $T \rightarrow R$ trajectory of ATCase. Main figure: E_{SWAXS} verus time; inset: trimer separation d_{sep} and relative rotation ϕ during that simulations, with positions every 5 ns labelled with diamonds. The decomposition shows that, while the rapid expansion during the first transition is triggered by restraints near $q = 1\,\mathrm{nm}^{-1}$, the subsequent rotation along ϕ was triggered by restraints at $q = 2\,\mathrm{nm}^{-1}$. (B) Partial SWAXS-driven MD trajectories of T and R ATCase, coupled to $0.2 \sim q \sim 1.6\,\mathrm{nm}^{-1}$ in order to fit only the first SAXS peak. Green lines indicate smoothed $R \rightarrow T$ trajectories, and red lines indicate smoothed $T \rightarrow R$ trajectories. Exclusion of the second peak during SWAXS-driven MD results in transitions primarily along d_{sep} with only diffusive rotation along ϕ. (C) Partial $R \rightarrow R$ and $R \rightarrow T$ SWAXS-driven trajectories coupled only to the second peak signal at $q = 2\,\mathrm{nm}^{-1}$. Red lines indicate $R \rightarrow R$ trajectories using $k_c = 500$, maroon lines indicate $R \rightarrow R$ trajectories using $k_c = 2500$, and green lines indicate $R \rightarrow T$ trajectories using $k_c = 500$. Those trajectories demonstrate that coupling purely to the SAXS signal at $q = 2\,\mathrm{nm}^{-1}$ results in a rotation along ϕ without any significant change in d_{sep}.

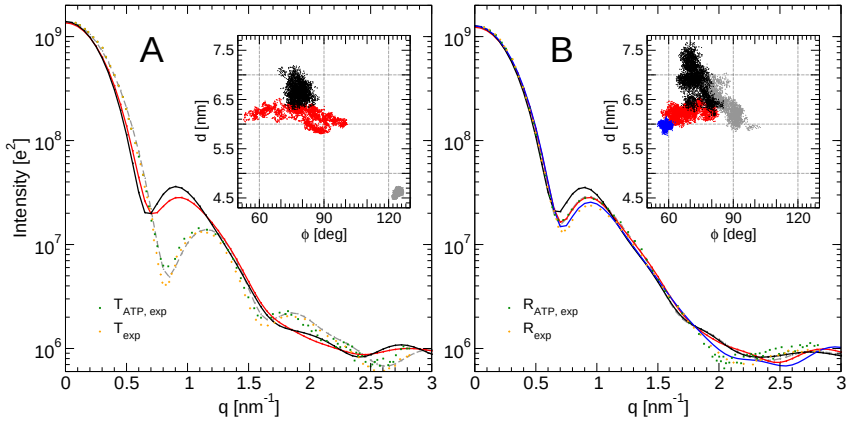

Figure S9. SAXS curves and trimer configuration of ATCase ensembles in free simulations before applying SWAXS-derived restraints (grey, dotted), during SWAXS-driven simulations (red and blue), and in free simulations after releasing the SWAXS-derived restraints (black). Good agreement with experimental R is found at low ϕ (blue line and dots in B). After releasing the SWAXS-derived restraints, the simulations partly returned to the solution R in free simulations. Insets show corresponding trajectories along the trimer separation d_{sep} and relative trimer rotation ϕ in their respective colors. (A) Trajectories starting from the T-state. (B) Trajectories starting from the R-state.

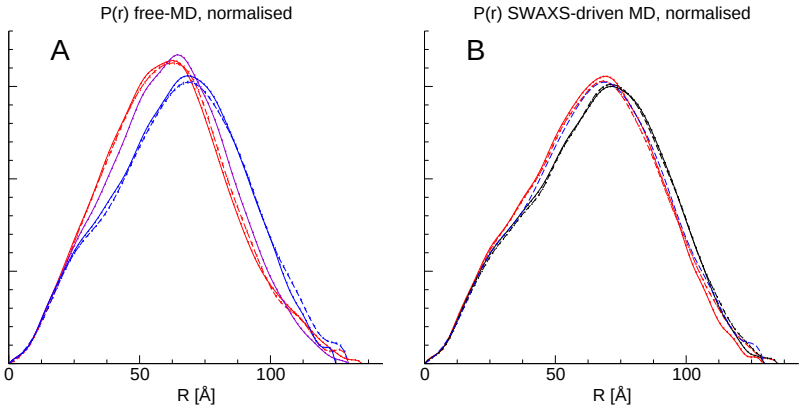

Figure S10. Pair distribution functions $P(r)$ for ATCase, calculated using GNOM based on curves in Figures 5A and S9. (A) $P(r)$ curves for ATCase. Color scheme as follows – Red-solid: T-ATCase unbiased simulations. Red-dashed: T-ATCase experiment. Purple-solid: R-ATCase unbiased position-restrained simulations. Blue-solid: R-ATCase unbiased simulations. Blue-dashed: R-ATCase experiment. (B) $P(r)$ curves for ATCase during and after SWAXS-driven MD. Color scheme as follows – Blue-dashed: Experimental R-ATCase, as in A. Red-dashed: T-ATCase driven to R-state spectra. Black-dashed: Relaxed T-ATCase after release of SWAXS constraints. Red-solid: R-ATCase driven to R-state spectra. Black-solid: Relaxed R-ATCase after release of SWAXS constraints.

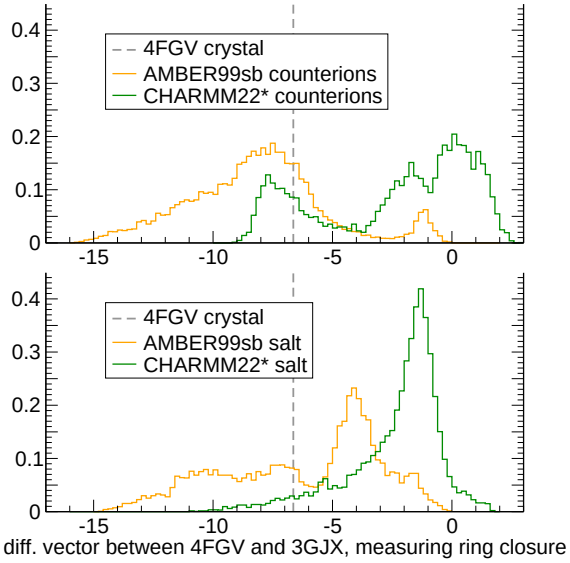

Figure S11. Conformations of CRM1 under different conditions, projected onto the difference vector between open (4FGV) and closed (3GJX) CRM1, following Monecke *et al.*(4). Location of 4FGV structure shown as grey lines. Top: Comparison of AMBER99sb and CHARMM22*, solvated only with counterions. Bottom: Comparison of AMBER99sb and CHARMM22*, solvated with 100 mM NaCl.

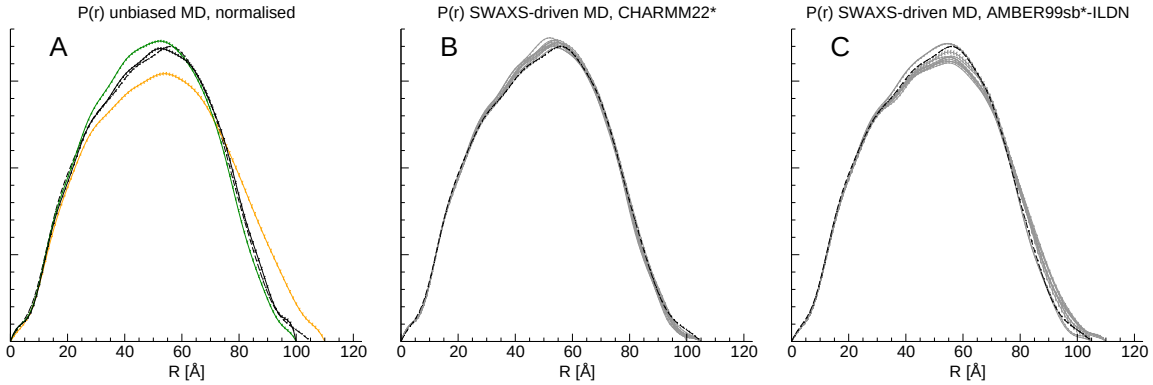

Figure S12. Pair distribution functions $P(r)$ for CRM1, calculated using GNOM based on curves in Figures 6B and S14A. (A) P(r) curves for ctCRM1. Color scheme as follows – green-solid: ctCRM1 unbiased simulations using CHARMM22* forcefield. orange-solid: ctCRM1 unbiased simulations using ABER99sb*-IDLN forcefield. black-dotted: ctCRM1 experiment. black-solid: Hybrid spectra of 45% free-ctCRM1 (ring-closed) and 55% restrained ctCRM1 (ring-open). (B) $P(r)$ curves for ctCRM1 during SWAXS-driven MD, under CHARMM22*. Color scheme as follows – black-dashed: target experimental ctCRM1 curve, as in A. grey-solid: individual SWAXS-driven replica. (C) $P(r)$ curves for ctCRM1 during SWAXS-driven MD, under AMBER99sb. Color scheme as follows – black-dashed: target experimental ctCRM1 curve, as in A. grey-solid: individual SWAXS-driven replica.

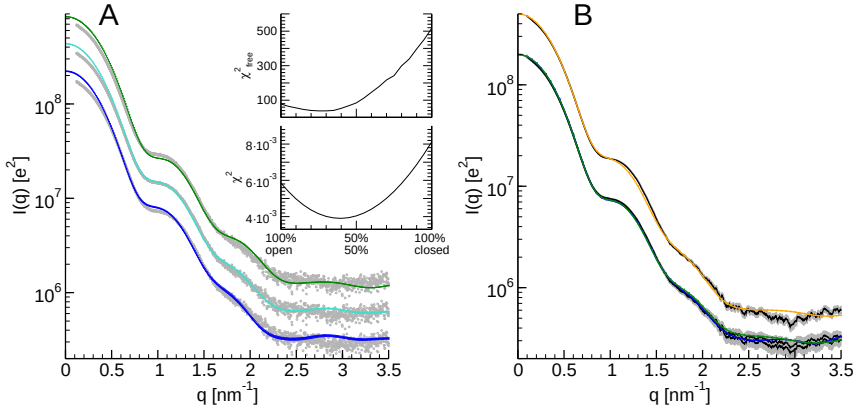

Figure S13. (A) SAXS curves of CRM1 ensembles under the CHARMM22* forcefield, in ring-open (blue) and ring-closed (green) conformations, and a linear combination of 60% open/40% closed spectra (cyan). The unsmoothed experimental spectra (circles) was fit to these curves and displayed for comparison, and curves were vertically shifted for clarity. Ensembles were generated from position-restrained and free simulations. In the insets, χ^2 fits were conducted on linear combinations of the open and closed curves according to Eq. 1 in the Supporting Material Methods, and χ^2_{free} fits were conducted according to Rambo and Tainer (25). The q-range for this fit is set between 0.5∼2.7 nm^{-1} to avoid the dominating influence of experimental noise at q∼0.3 nm^{-1}. (B) SAXS patterns of CRM1 ensembles after SWAXS-driven MD. Blue and green curves: CHARMM22* ensembles. Orange: AMBER99sb ensemble. Black: smoothed experimental curve with errors (grey).

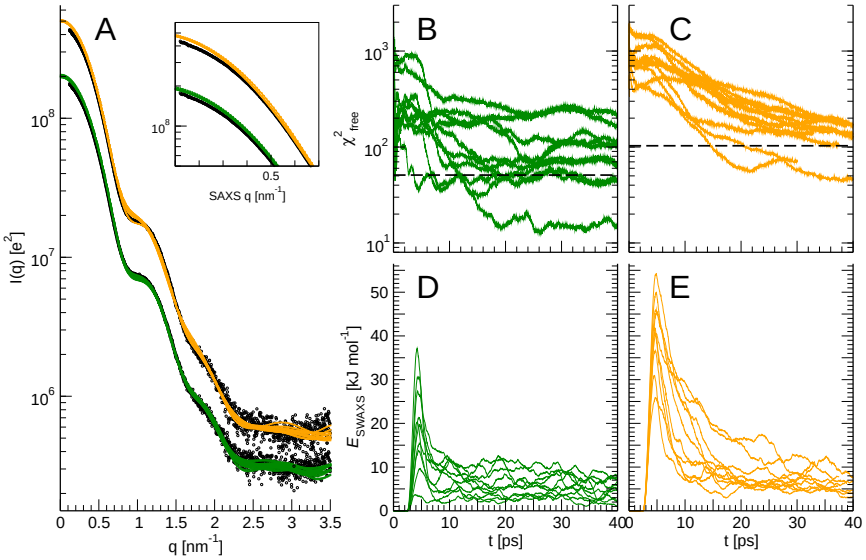

Figure S14. χ^2_{free} analysis of ctCRM1 SWAXS-driven trajectories using the CHARMM22* (green) or the AMBER99sb force field (orange). (A) Exponentially averaged calculated SAXS curves of CRM1, at the final simulation timepoint of SWAXS-driven MD. All replica are plotted as thin lines over the fitted, unsmoothed experimental data (circles). AMBER99sb curves are vertically offset for clarity. A magnification of the SAXS region is shown on the inset. (B/C) χ^2_{free} values during SWAXS-driven MD. Values are high due to significant underfitting at the Shannon bins between $0.5 \sim 0.7$ nm^{-1}, and around 1 nm^{-1}, suggesting that the MD forcefield imposed tight restraints on possible protein conformations. Thick dotted lines represent χ^2_{free} of the ensemble SAXS curve in Fig. S13B after averaging over all ten replica. (D/E) E_{SWAXS} values during SWAXS-driven MD, demonstrating that E_{SWAXS} provides similar information as χ^2_{free}.