

# Summary of PTE B-Factor analysis & Ensemble Refinement

## B-FACTOR ANALYSIS:

[B-Factor analysis of S Traj](#)

[B-factor analysis of R Traj](#)

## ENSEMBLE REFINEMENT:

[Ensemble Refinement of S Traj](#)

[Ensemble Refinement of R Traj](#)

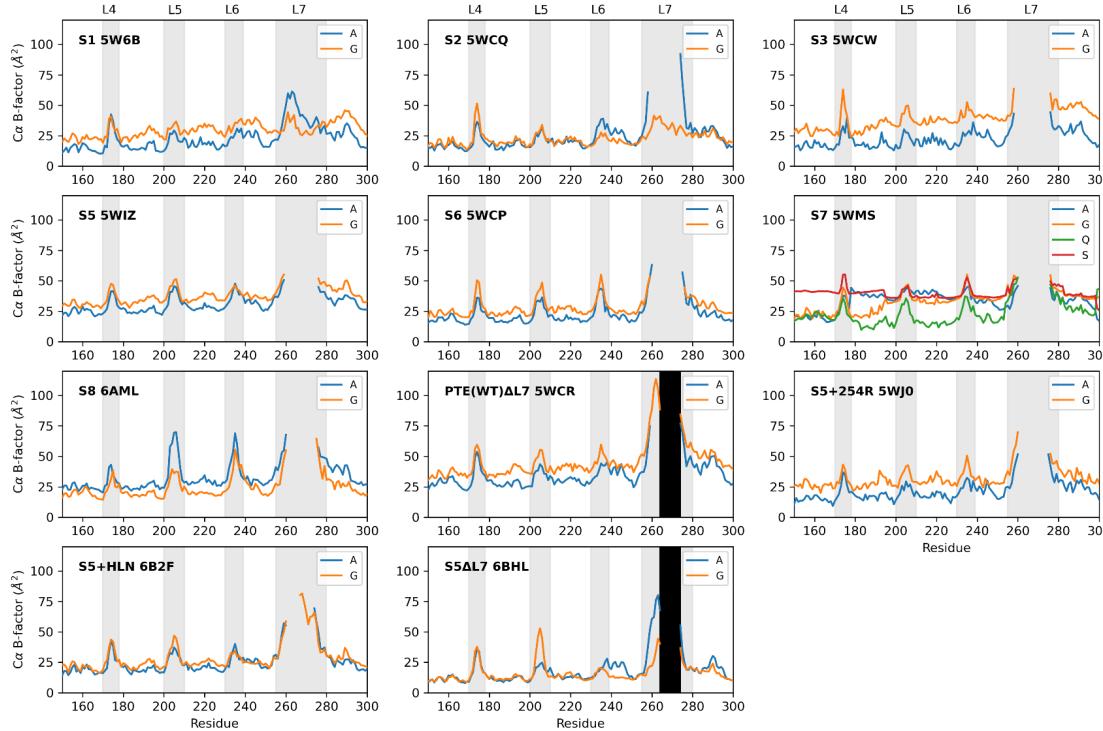
[Ensemble Refinement Tables](#)

## METHODS

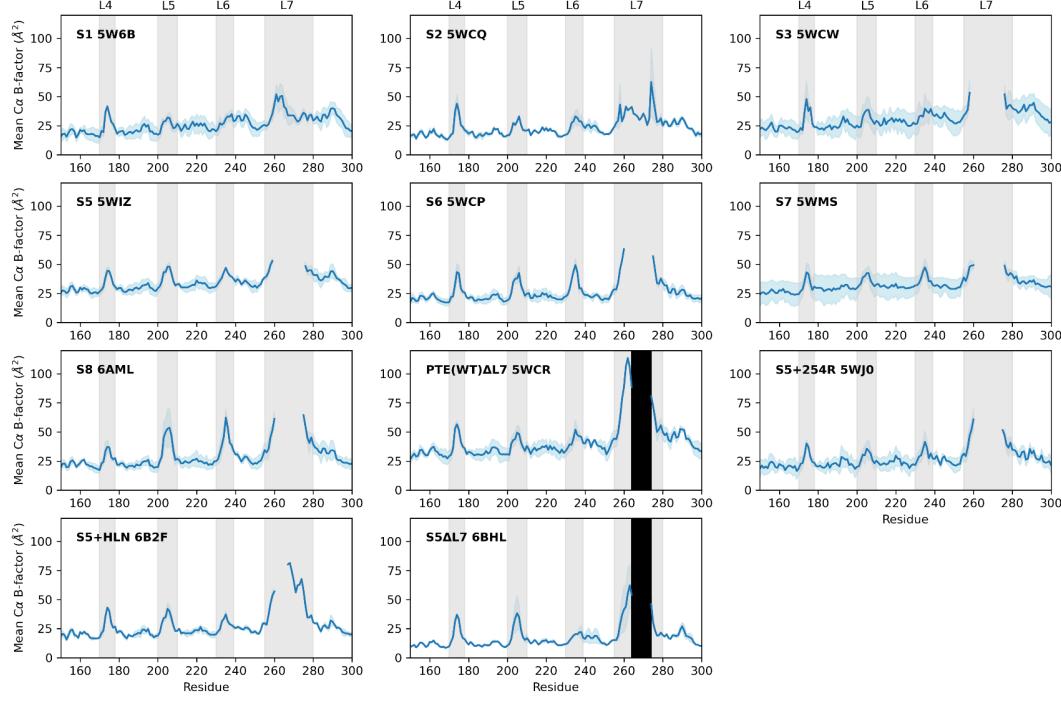
## REFERENCES

## B-FACTOR ANALYSIS:

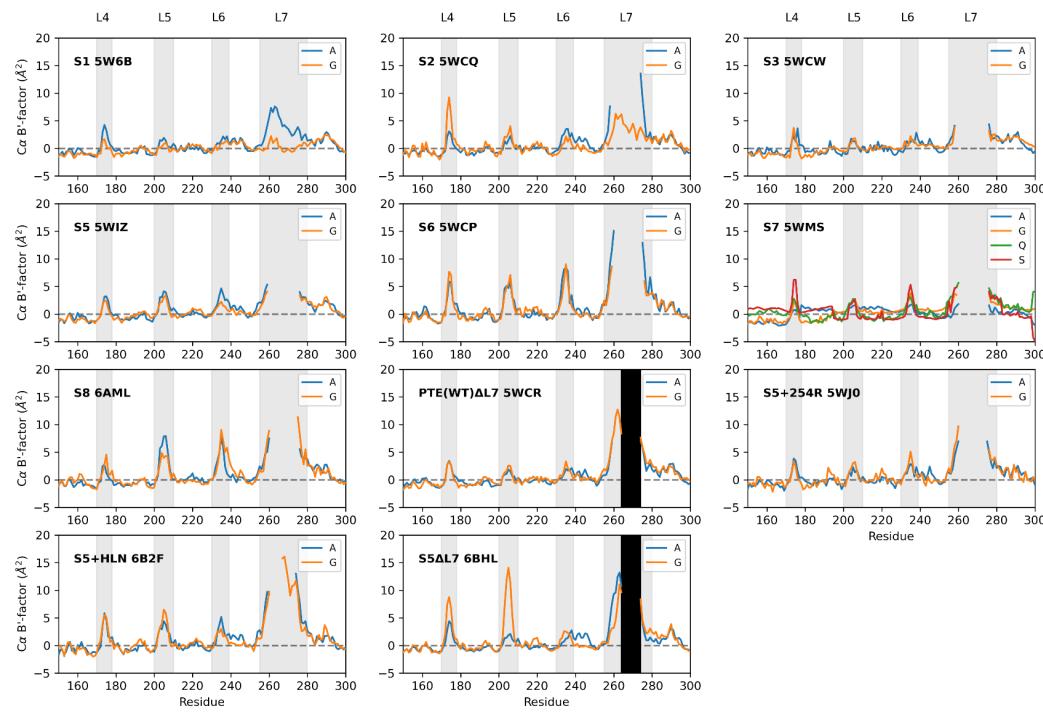
### B-Factor analysis of S Traj



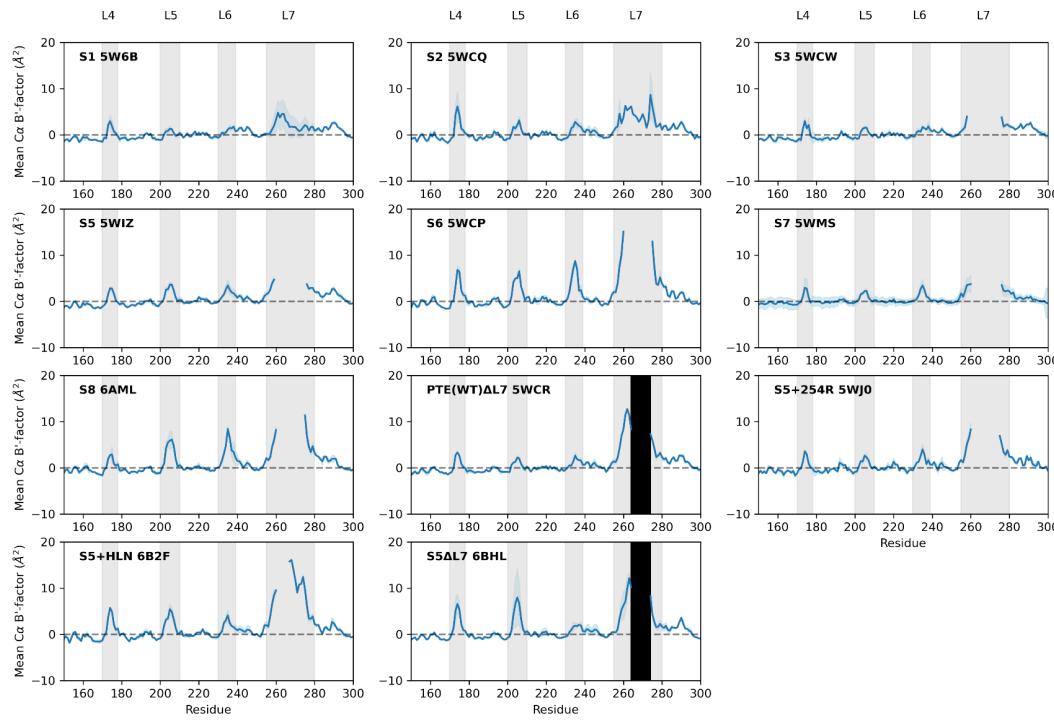
**Supplementary Figure X. Alpha-carbon B-factor analysis.** Alpha-carbon isotropic B-factors for residues 150–300 of each crystal structure described in this study. Each chain has been analyzed separately and colored differently (black, blue, orange, green for subsequent chains in each PDB). Where alternate conformers were modelled in the PDB, conformer B was omitted from analysis. Key loops L4 (residues 170–178), L5 (residues 200–210), L6 (residues 230–239), and L7 (residues 255–280) are shaded in grey. Loop deletion shaded black.



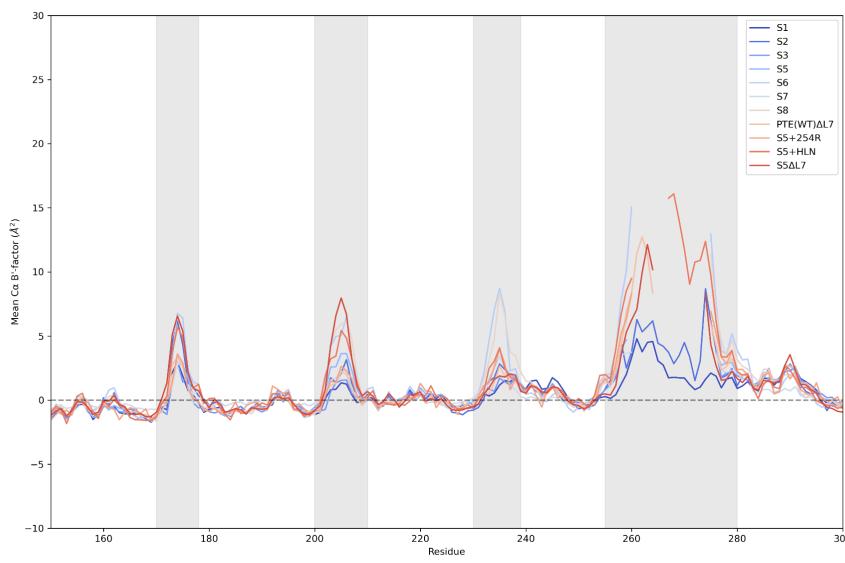
**Supplementary Figure X. Mean alpha-carbon B-factor analysis.** Mean (+SD) alpha-carbon isotropic B-factors for residues 150-300 of each crystal structure described in this study. Where alternate conformers were modelled in the PDB, conformer B was omitted from analysis. Key loops L4 (residues 170–178), L5 (residues 200–210), L6 (residues 230–239), and L7 (residues 255–280) are shaded in grey. Loop deletion shaded black.



**Supplementary Figure X. Base-line normalized alpha-carbon B-factor (B'-factor) analysis of the crystal structures presented in this work.**  
 Normalized alpha-carbon isotropic B-factors (B'-factor) for residues 150-300 of each crystal structure in this study. Each chain is analyzed separately and colored differently (black, blue, orange, green for subsequent chains in each PDB). Where alternate conformers were modelled in the PDB, conformer B was omitted from analysis. Key loops L4 (residues 170–178), L5 (residues 200–210), L6 (residues 230–239), and L7 (residues 255–280) are shaded in grey.

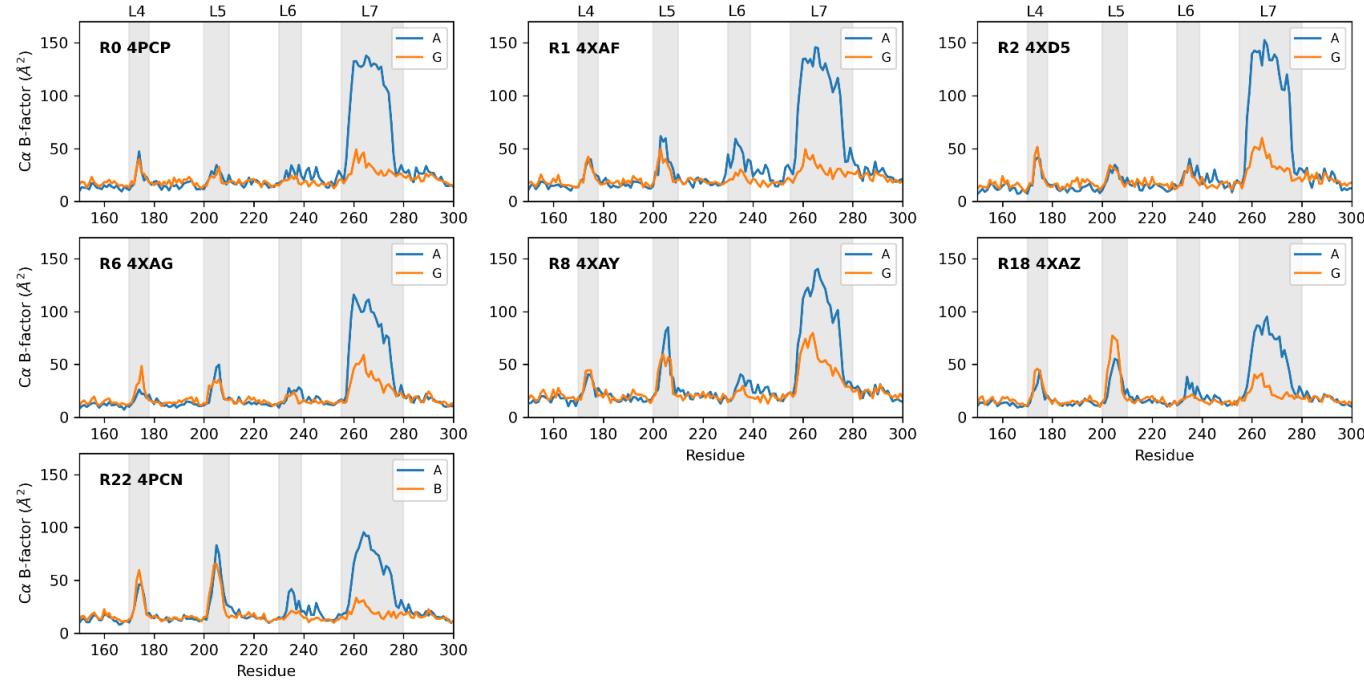


**Supplementary Figure X. Mean base-line normalized alpha-carbon B-factor (B'-factor) analysis of the crystal structures presented in this work.** Normalized alpha-carbon isotropic B-factors (B'-factor) for residues 150–300 of each crystal structure in this study. Each chain is analyzed separately and colored differently (black, blue, orange, green for subsequent chains in each PDB). Where alternate conformers were modelled in the PDB, conformer B was omitted from analysis. Key loops L4 (residues 170–178), L5 (residues 200–210), L6 (residues 230–239), and L7 (residues 255–280) are shaded in grey.

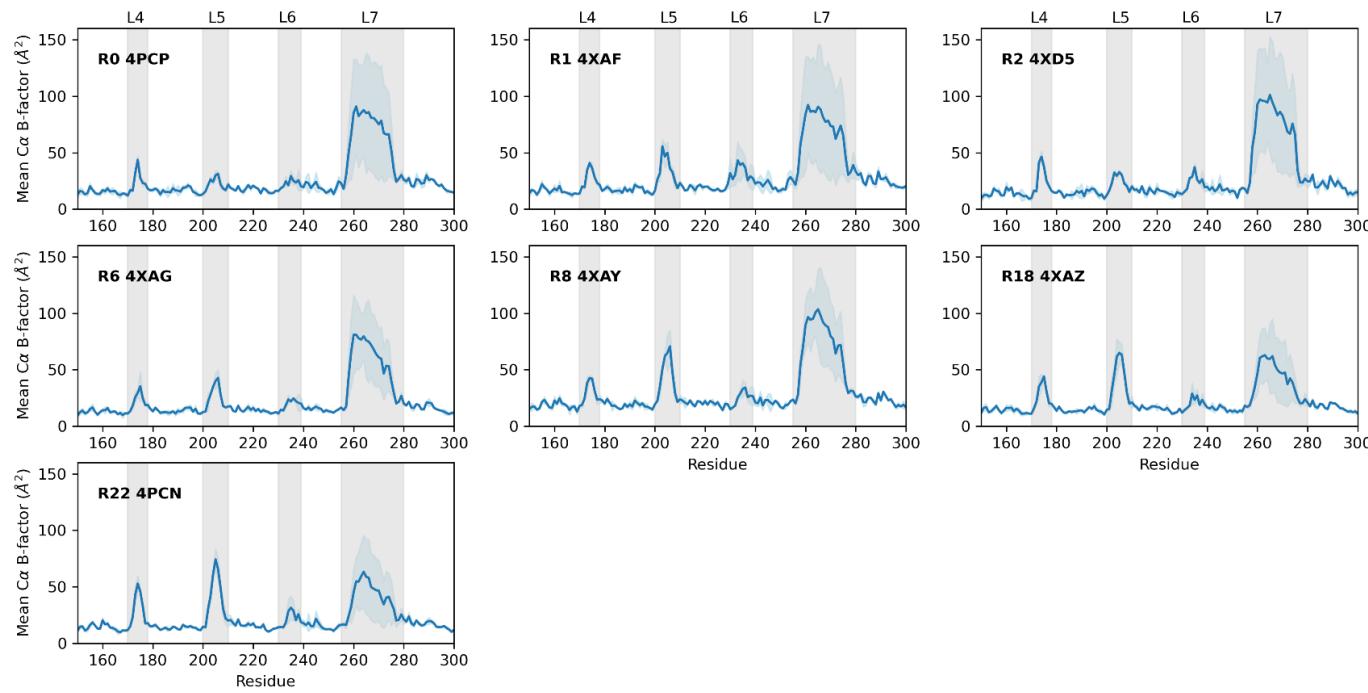


**Supplementary Figure X. Mean base-line normalized alpha-carbon B'-factor analysis of the crystal structures presented in this work (averaged across all chains in structure).** Mean normalized alpha-carbon isotropic B-factors (B'-factor) for residues 150-300 of S1-S8. Data represents the mean B'-factor when averaged across each chain in the crystal structure. Where alternate conformers were modelled in the PDB, conformer B was omitted from analysis. Key loops L4 (residues 170–178), L5 (residues 200–210), L6 (residues 230–239), and L7 (residues 255–280) are shaded in grey.

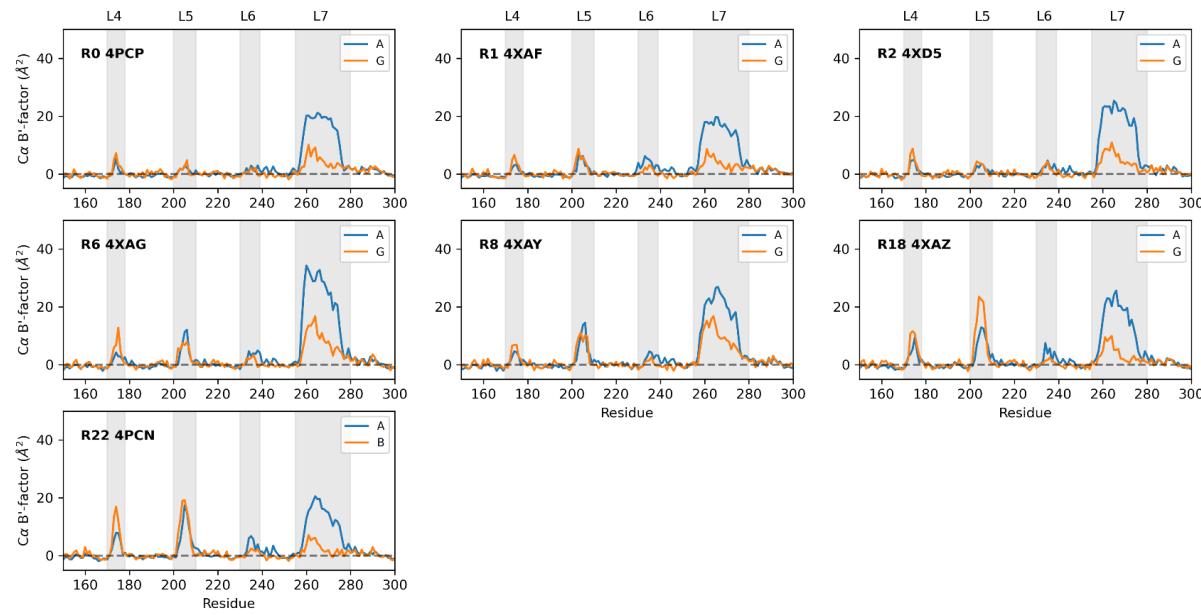
## B-factor analysis of R Traj



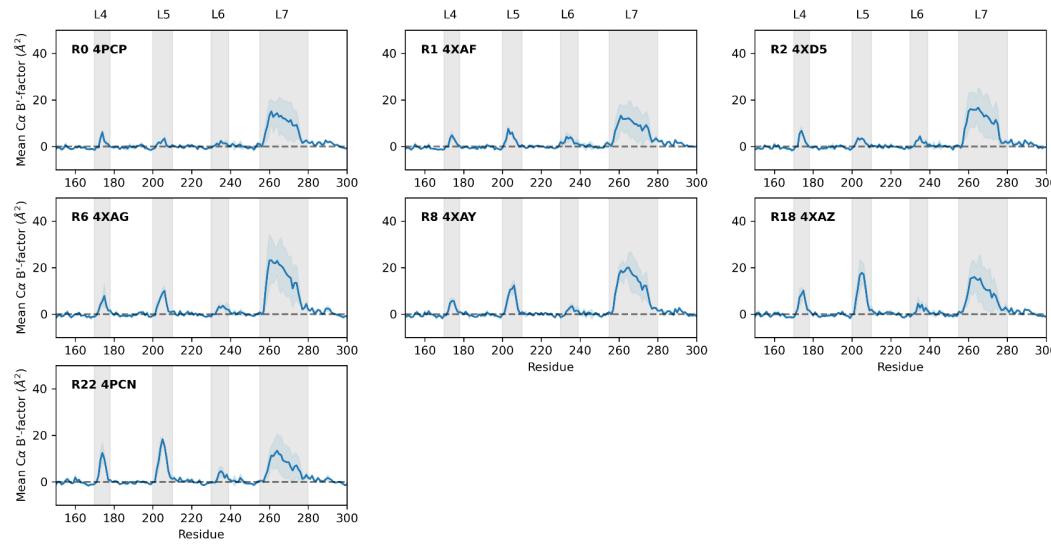
**Supplementary Figure X. Alpha-carbon B-factor analysis.** Alpha-carbon isotropic B-factors for residues 150–300 of each crystal structures on the Arg trajectory. Each chain has been analyzed separately and colored differently (black, blue, orange, green for subsequent chains in each PDB). Where alternate conformers were modelled in the PDB, conformer B was omitted from analysis. Key loops L4 (residues 170–178), L5 (residues 200–210), L6 (residues 230–239), and L7 (residues 255–280) are shaded in grey. Loop deletion shaded black.



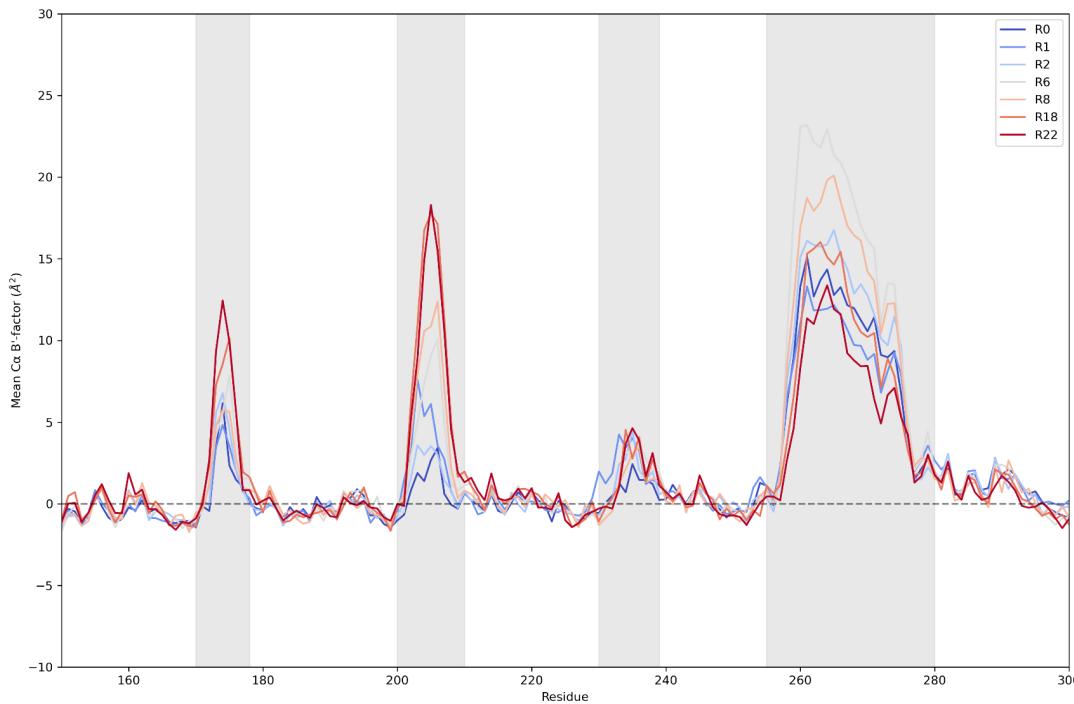
**Supplementary Figure X. Mean alpha-carbon B-factor analysis.** Mean (+SD) alpha-carbon isotropic B-factors for residues 150-300 of each crystal structure along the Arg trajectory. Where alternate conformers were modelled in the PDB, conformer B was omitted from analysis. Key loops L4 (residues 170–178), L5 (residues 200–210), L6 (residues 230–239), and L7 (residues 255–280) are shaded in grey.



**Supplementary Figure X. Base-line normalized alpha-carbon B-factor (B'-factor) analysis of the Arg trajectory.** Normalized alpha-carbon isotropic B-factors (B'-factor) for residues 150-300 of each crystal structure along the Arg trajectory. Each chain is analyzed separately and colored differently (black, blue, orange, green for subsequent chains in each PDB). Where alternate conformers were modelled in the PDB, conformer B was omitted from analysis. Key loops L4 (residues 170–178), L5 (residues 200–210), L6 (residues 230–239), and L7 (residues 255–280) are shaded in grey.



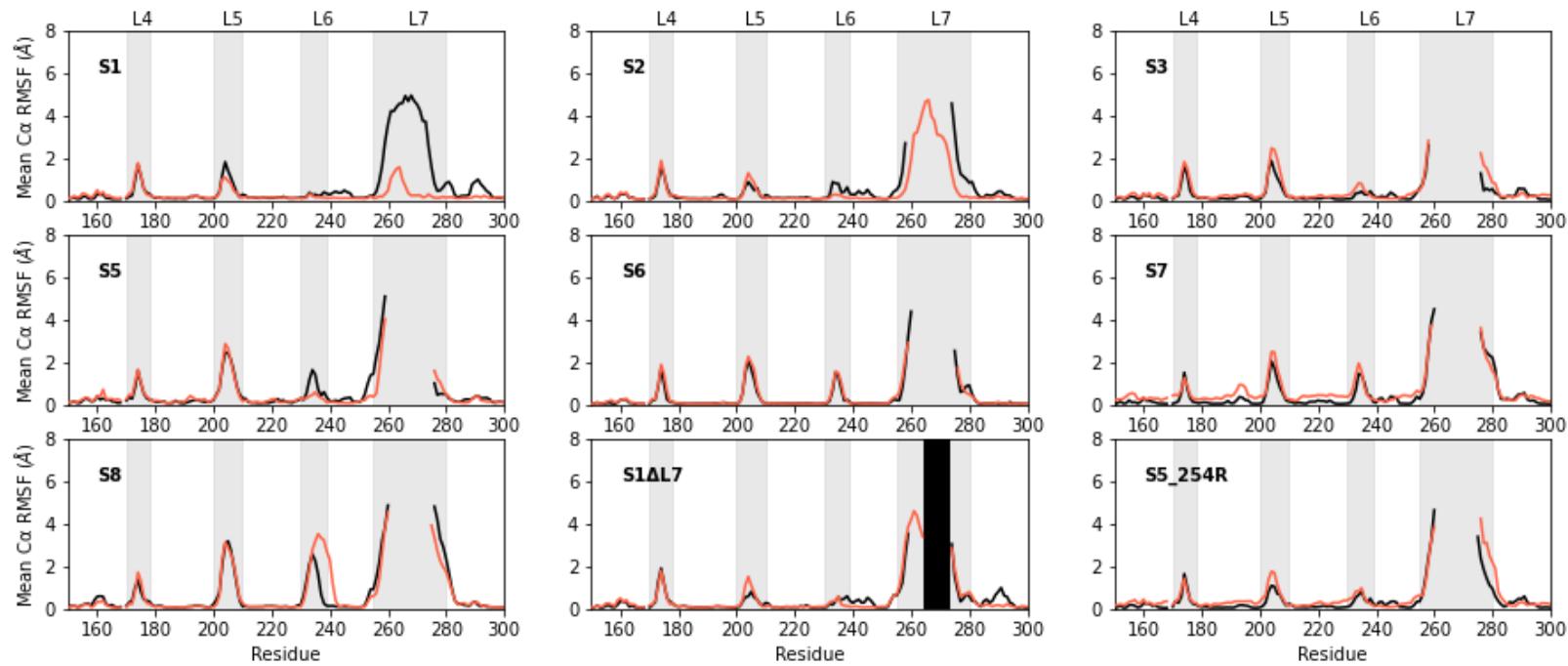
**Supplementary Figure X. Mean base-line normalized alpha-carbon B-factor (B'-factor) analysis of the Arg trajectory.** Normalized alpha-carbon isotropic B-factors (B'-factor) for residues 150-300 of each crystal structure in this study. Each chain is analyzed separately and colored differently (black, blue, orange, green for subsequent chains in each PDB). Where alternate conformers were modelled in the PDB, conformer B was omitted from analysis. Key loops L4 (residues 170–178), L5 (residues 200–210), L6 (residues 230–239), and L7 (residues 255–280) are shaded in grey.



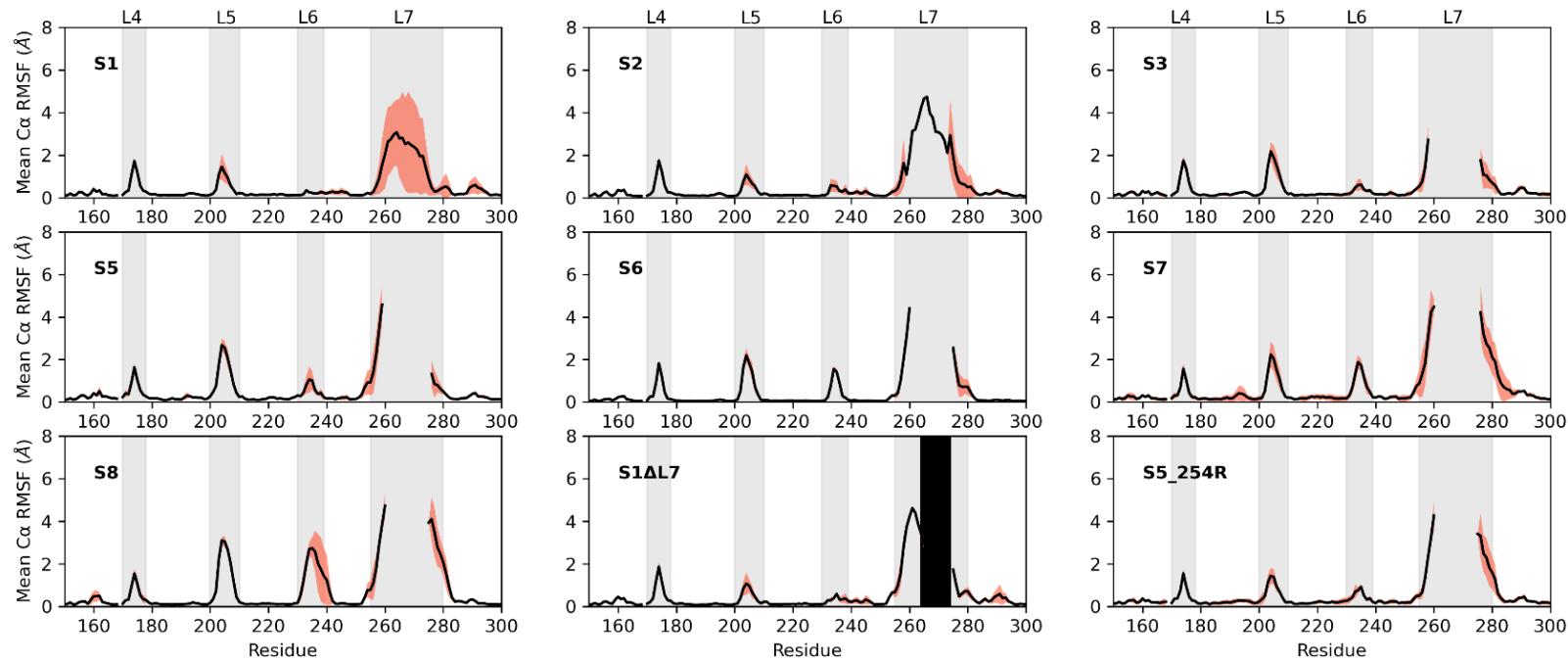
**Supplementary Figure X. Mean base-line normalized alpha-carbon B'-factor analysis of the crystal structures presented in this work (averaged across all chains in structure).** Mean normalized alpha-carbon isotropic B-factors (B'-factor) for residues 150-300 of S1-S8. Data represents the mean B'-factor when averaged across each chain in the crystal structure. Where alternate conformers were modelled in the PDB, conformer B was omitted from analysis. Key loops L4 (residues 170–178), L5 (residues 200–210), L6 (residues 230–239), and L7 (residues 255–280) are shaded in grey.

ENSEMBLE REFINEMENT:

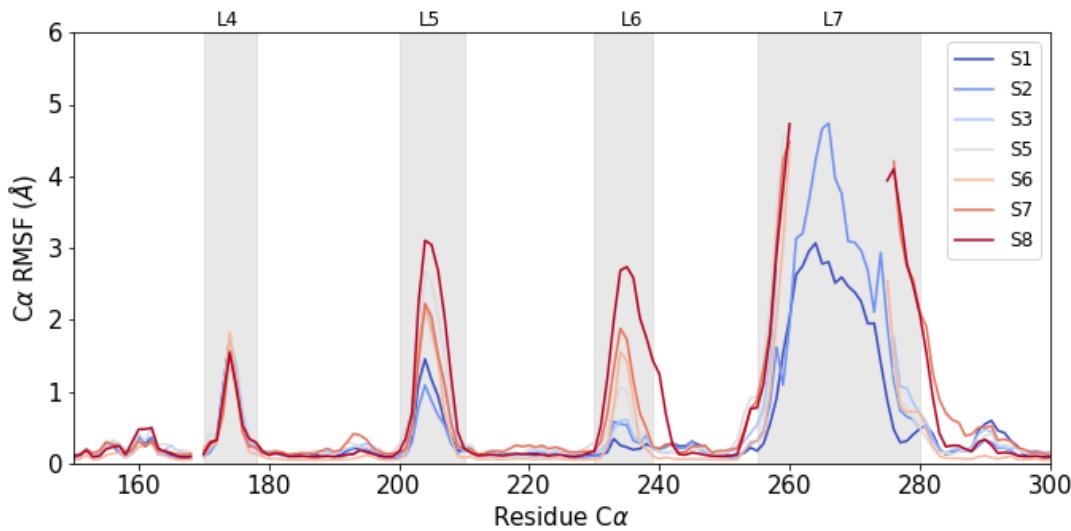
## Ensemble Refinement of S Traj



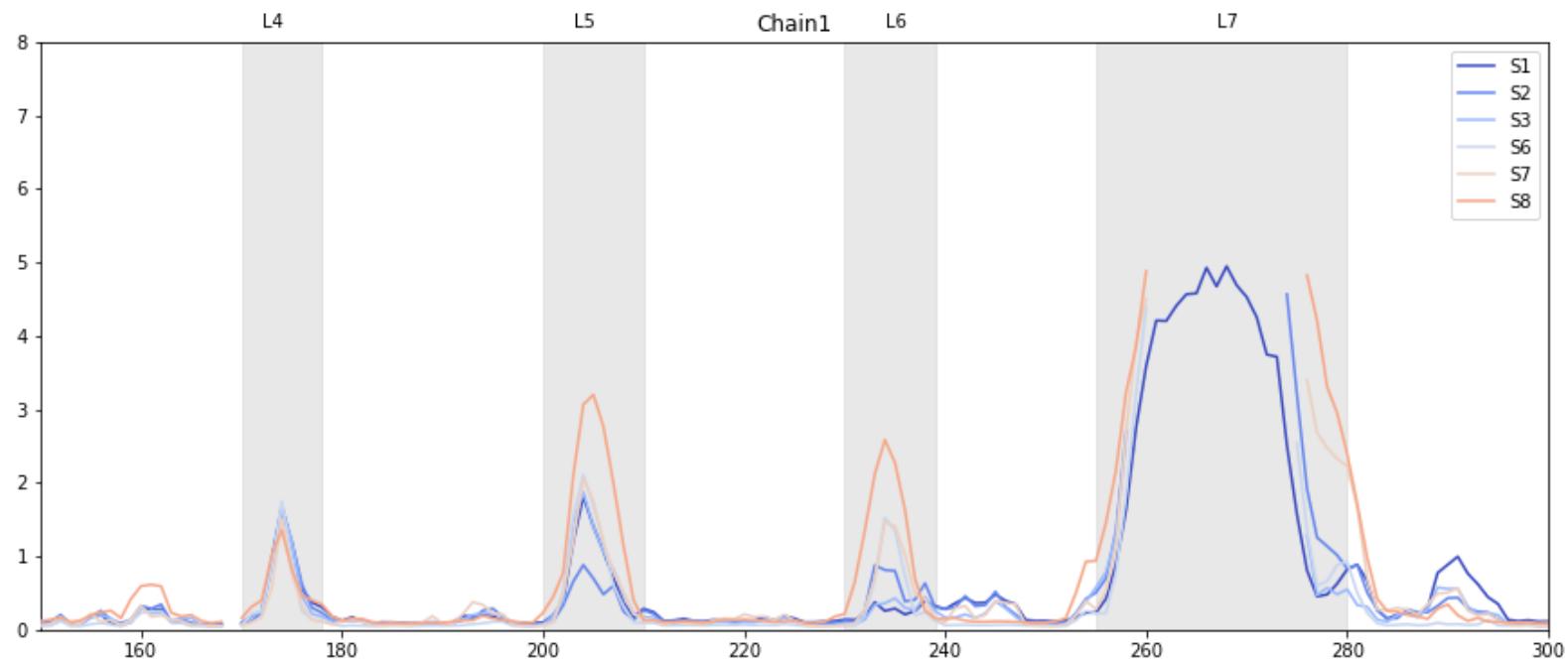
**Supplementary Figure X. a** Ca RMSF (residues 150-300) plots from ensemble refinement of the Ser trajectory structures. For each replicate, the mean RMSF of each residue was calculated for first and second chains (i.e. RMSFs were averaged between chains in the structure). Plots show the mean ( $n=5$ ) for chain 1 (cyan) and chain 2 (magenta) across the top 5 replicates.



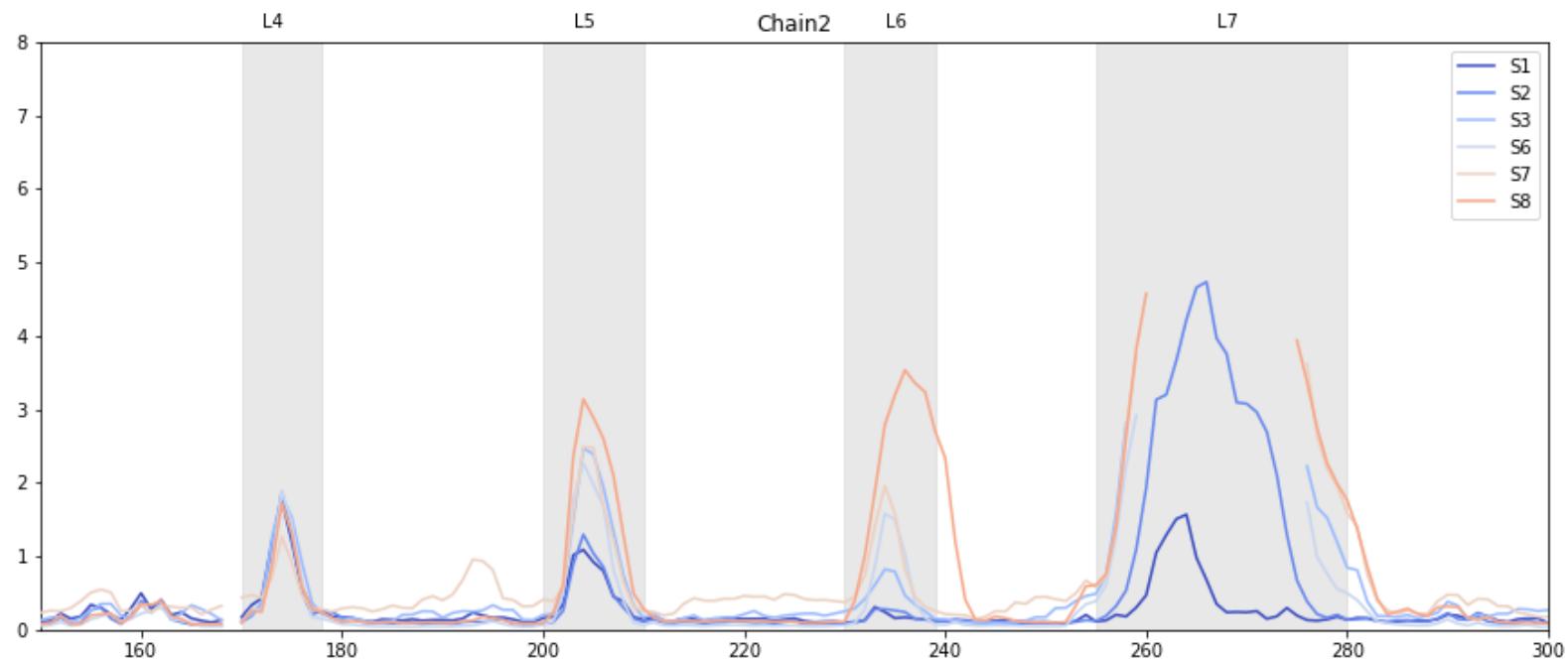
**Supplementary Figure X. a**  $\text{C}\alpha$  RMSF (residues 150-300) plots from ensemble refinement of the Ser trajectory structures. For each replicate, the mean RMSF of each residue was calculated (i.e. RMSFs were averaged between chains in the structure). Plots show the mean  $\pm$  standard deviation of these mean RMSFs (shaded area,  $n=5$ ).



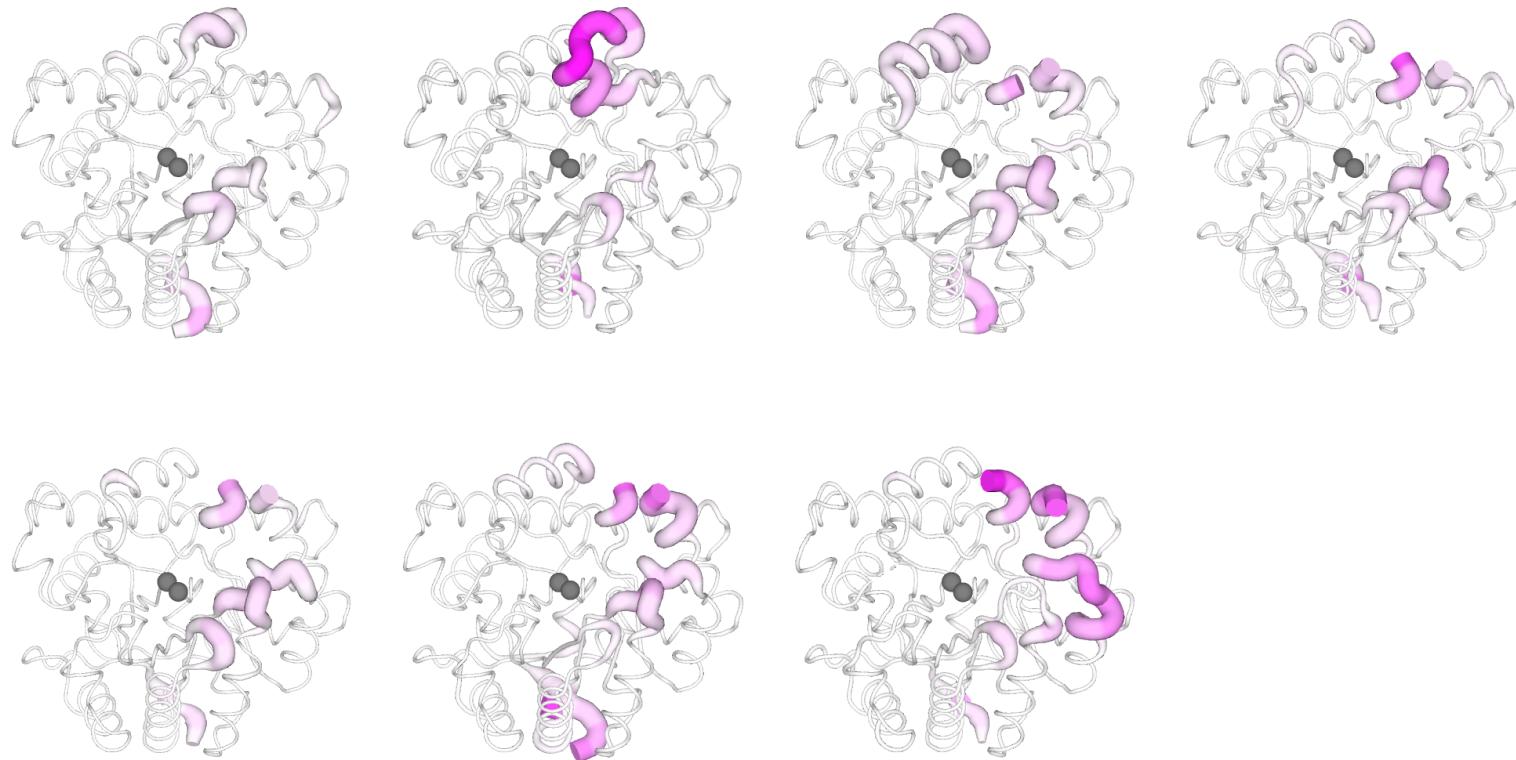
**Supplementary Figure X.** C $\alpha$  RMSF (residues 150-300) from ensemble refinement of the Ser trajectory structures. Plots are means per-residue RMSFs calculated by taking the averaging across all chains in the structure, then averaging the RMSFs of the top five replicates.



Only considering chain 1 (normally chain A).



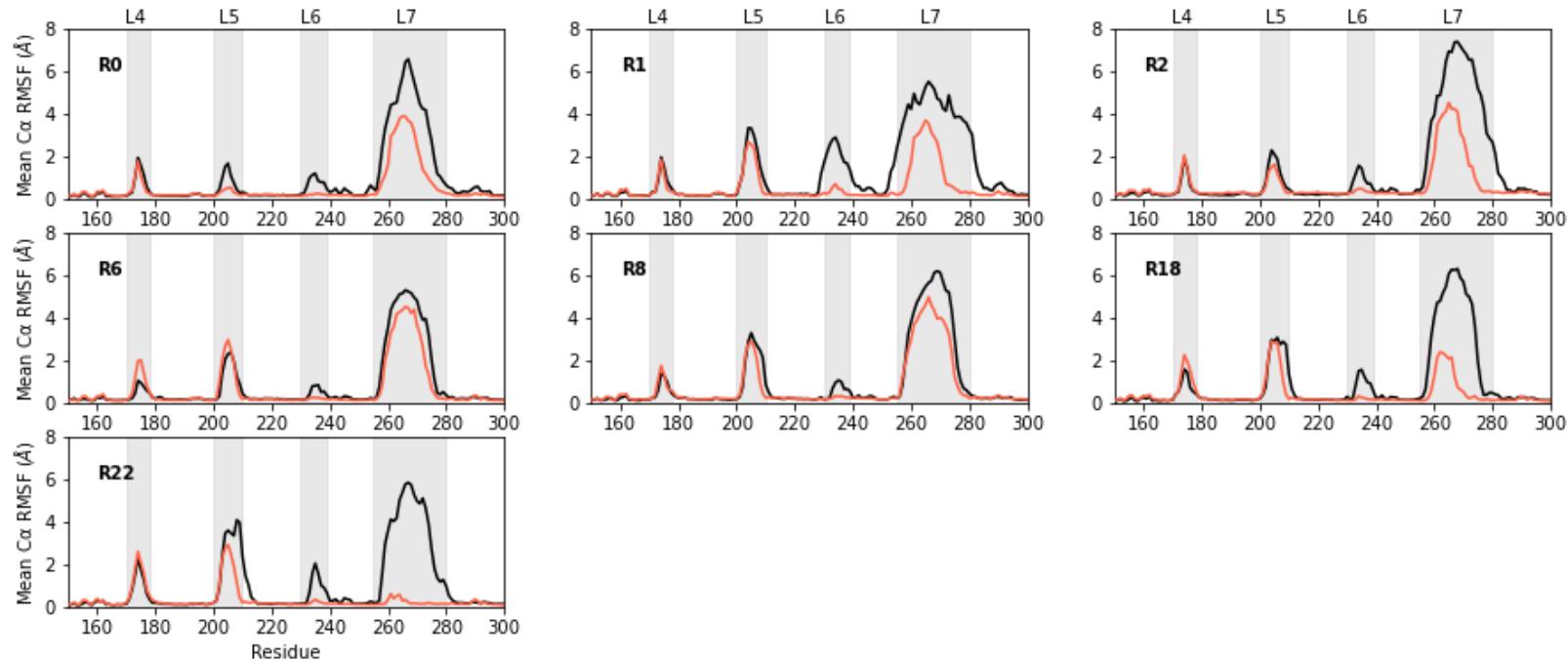
Only considering chain 2 (normally chain G). (most pronounced changes) - use this for pymol figures?



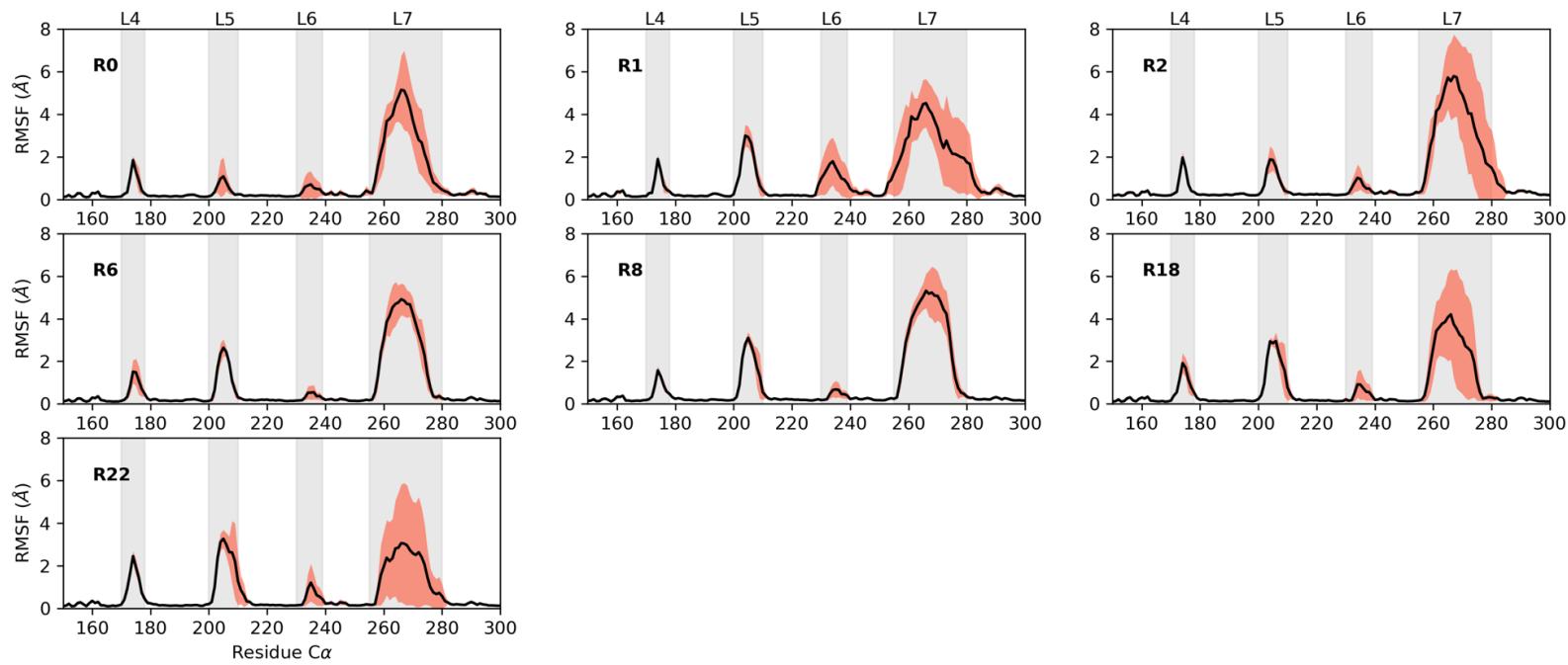
Pymol figure showing changes in loop RMSF of chain G along the S trajectory. In each case, show as putty representation using RMSF values rather than B-factor but based on crystal structure position of L7. S1, S2, S3, S5, S6, S7, S8.

RMSF for chain G only. - (set cartoon\_putty\_transform, 0); spectrum b, white magenta. minimum=0. maximum=2000; atom.set\_bfactor((8 \* math.pi\*\*2) \* (rmsf\*\*2)\*100) (i.e. **includes scaling factor!**)

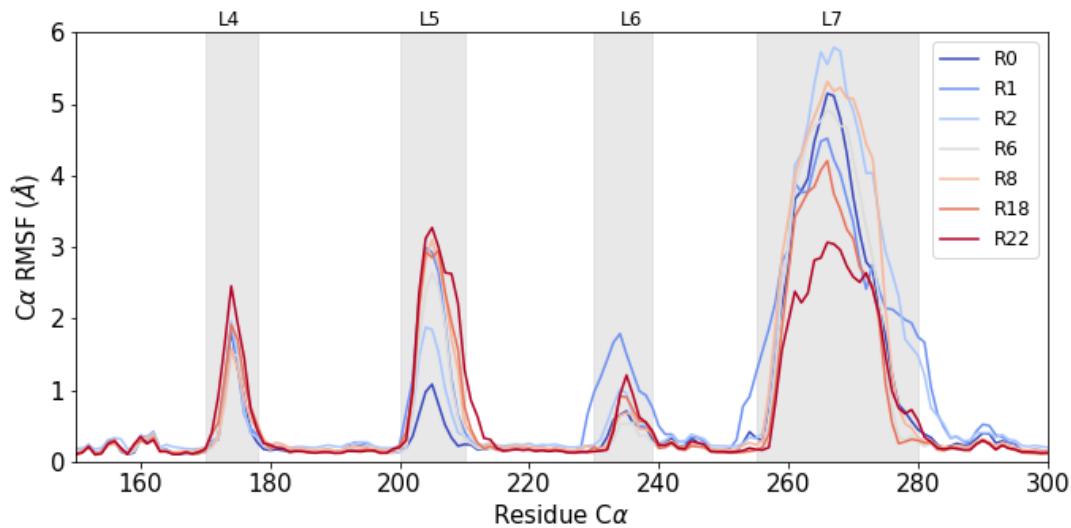
## Ensemble Refinement of R Traj



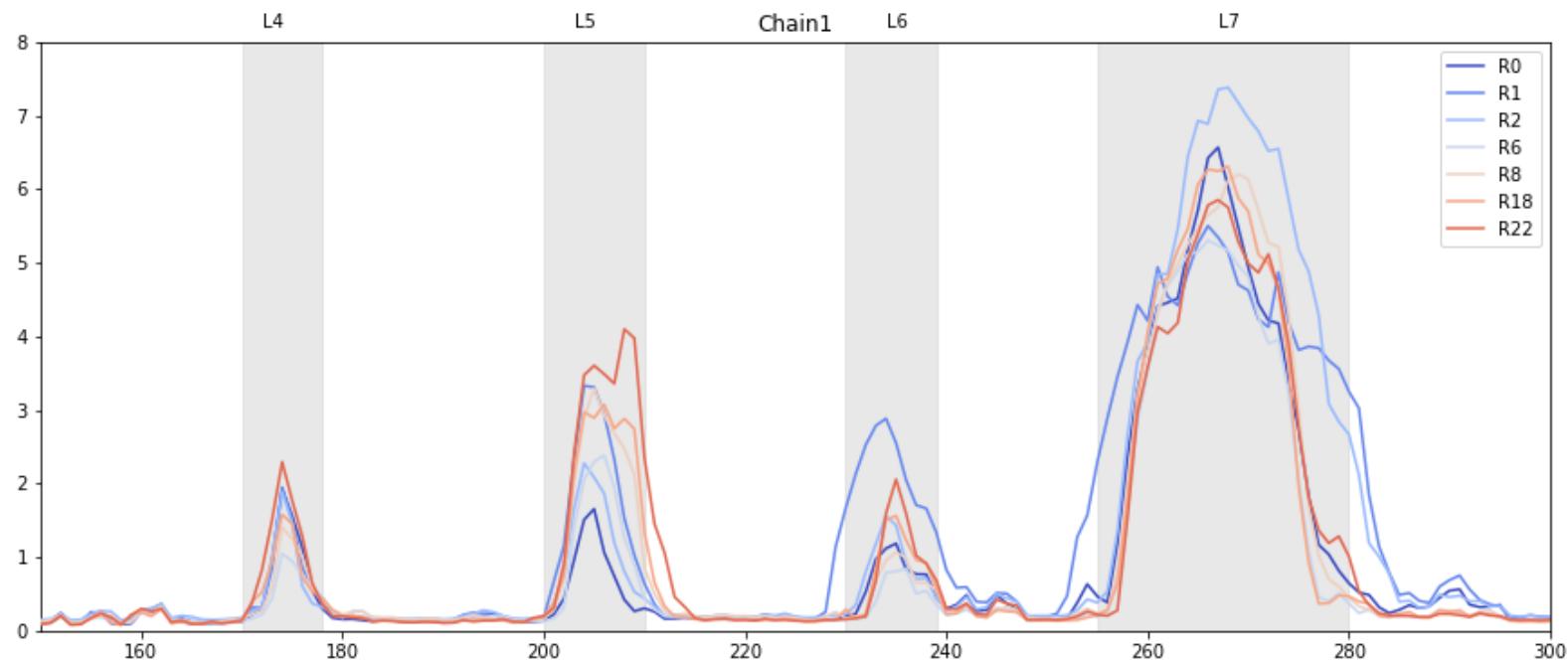
**Supplementary Figure X. a** Ca RMSF (residues 150-300) plots from ensemble refinement of the Ser trajectory structures. For each replicate, the mean RMSF of each residue was calculated (i.e. RMSFs were averaged between chains in the structure). Plots show the mean  $\pm$  standard deviation of these mean RMSFs (shaded area, n=5). The two first chains in pdb analysed separately (typically chain A/G), as shown in black, tomato colors.



**Supplementary Figure X. a**  $\text{Ca}$  RMSF (residues 150-300) plots from ensemble refinement of the Ser trajectory structures. For each replicate, the mean RMSF of each residue was calculated (i.e. RMSFs were averaged between chains in the structure). Plots show the mean  $\pm$  standard deviation of these mean RMSFs (shaded area,  $n=5$ ).

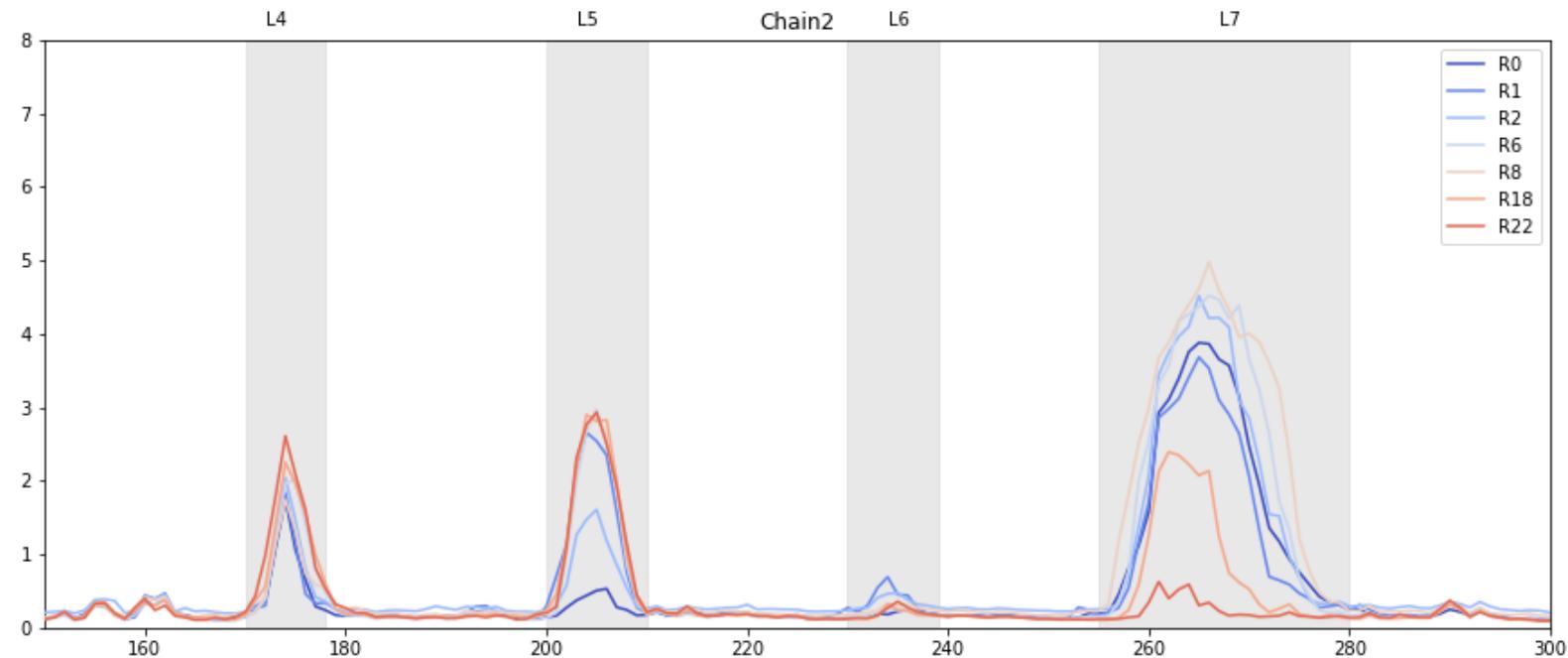


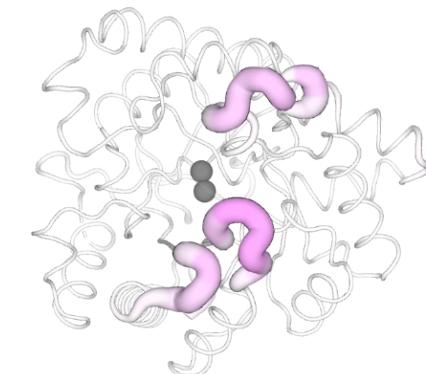
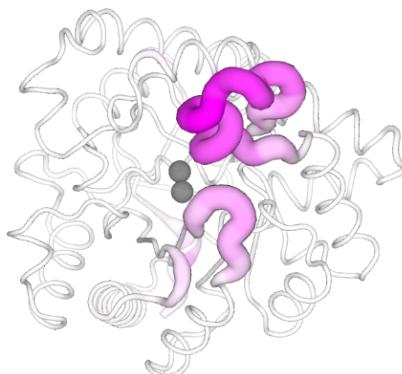
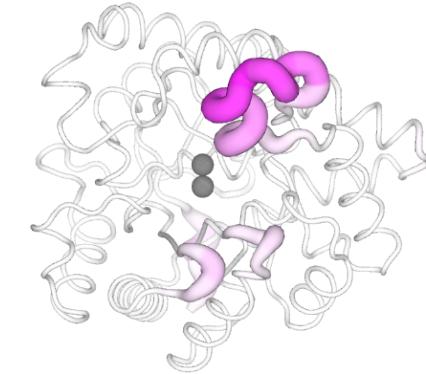
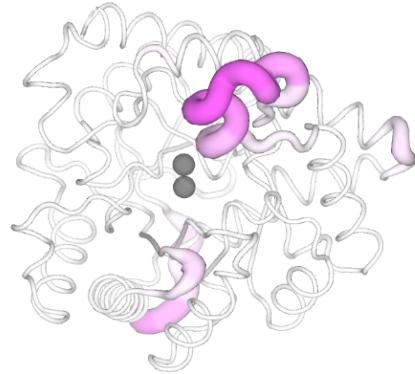
**Supplementary Figure X.** C $\alpha$  RMSF (residues 150-300) from ensemble refinement of the Arg trajectory structures. Plots are means per-residue RMSFs calculated by taking the averaging across all chains in the structure, then averaging the RMSFs of the top five replicates.



Chain 1 trajectory. Less changes in L7.

Chain 2 trajectory - most pronounced changes with decrease in L7 RMSF, and increase in L5. - Use this for pymol figure.





Pymol figure showing changes in loop RMSF of chain G along the R trajectory. R0, R1, R2, R6, R8, R18, R22.

RMSF for chain G only. - (set cartoon\_putty\_transform, 0); spectrum b, white magenta. minimum=0. maximum=2000; atom.set\_bfactor((8 \* math.pi\*\*2) \* (rmsf\*\*2)\*100). (i.e. **includes scaling factor!**)



## Ensemble Refinement Tables

**Supplementary Table X.** Statistics for ensemble refinement of structures along the Ser trajectory.

	Dataset								
	S1	S2	S3	S5	S6	S7	S8	R0deltaL7	S5+254R
PDB ID	5W6B	5WCQ	5WCW	5WIZ	5WCP	5WMS	6AML	5WCR	5WJ0
p <sub>TLS</sub> (%)	0.9	0.9	0.8	0.9	1	0.9	0.9	0.9	0.9
Wxray (K)	10	5	5	5	10	5	10	10	2.5
$\tau_x$ (ps)	1	1	0.6	1	0.6	1	1	1	1
No. of models*	68.8 ± 6.4 (72)	77.4 ± 8.6 (84)	56.0 ± 8.5 (60)	50.8 ± 5.2 (50)	65.6 ± 2.8 (67)	51.8 ± 8.4 (42)	84.0 ± 0.0 (84)	56.4 ± 5.8 (50)	59.0 ± 5.3 (63)
R <sub>free</sub> *	0.1714 ± 0.0007 (0.1703)	0.1750 ± 0.0015 (0.1728)	0.2133 ± 0.0006 (0.2123)	0.2111 ± 0.0008 (0.2096)	0.1634 ± 0.0006 (0.1621)	0.2811 ± 0.0005 (0.2804)	0.1643 ± 0.0020 (0.1608)	0.1822 ± 0.0009 (0.1818)	0.1984 ± 0.0009 (0.1969)
R <sub>work</sub> *	0.1290 ± 0.0004 (0.1284)	0.1380 ± 0.0007 (0.1367)	0.1770 ± 0.0007 (0.1759)	0.1621 ± 0.0003 (0.1622)	0.1362 ± 0.0005 (0.1354)	0.2357 ± 0.0011 (0.2374)	0.1388 ± 0.0015 (0.1362)	0.1491 ± 0.0005 (0.1492)	0.1624 ± 0.0010 (1613)

\* Data is mean ± standard deviation for five random seed repeats that gave the lowest R<sub>free</sub> out of 10 random seed repeats. Values in parentheses correspond to the ensemble with the lowest R<sub>free</sub>.

**Supplementary Table X.** Statistics for ensemble refinement of structures along the Arg trajectory.

	Dataset							
	R0	R1	R2	R6	R8	R18	R22	
PDB ID	4PCP	4XAF	4XD5	4XAG	4XAY	4XAZ	4PCN	
p <sub>TLS</sub> (%)	0.9	0.8	0.8	0.9	0.9	0.9	0.9	
Wxray (K)	5	2.5	2.5	2.5	2.5	5	2.5	
$\tau_x$ (ps)	1	0.6	0.3	0.6	1	1	1	
No. of models*	76 ± 11 (84)	70.2 ± 3.9 (67)	47.2 ± 3.4 (50)	66.8 ± 9.1 (75)	73.2 ± 9.4 (84)	79.2 ± 14.6 (100)	77.4 ± 8.6 (72)	
R <sub>free</sub> *	0.1713 ± 0.0018 (0.1695)	0.1858 ± 0.0010 (0.1842)	0.1781 ± 0.0010 (0.1770)	0.1715 ± 0.0011 (0.1695)	0.1829 ± 0.0010 (0.1812)	0.1633 ± 0.0008 (0.1620)	0.1646 ± 0.0007 (0.1634)	
R <sub>work</sub> *	0.1313 ± 0.0018 (0.1294)	0.1442 ± 0.0013 (0.1434)	0.1331 ± 0.0007 (0.1329)	0.1336 ± 0.0012 (0.1315)	0.1379 ± 0.0007 (0.1371)	0.1339 ± 0.0010 (0.1327)	0.1293 ± 0.0007 (0.1293)	

\* Data is mean ± standard deviation for five random seed repeats that gave the lowest R<sub>free</sub> out of 10 random seed repeats. Values in parentheses correspond to the ensemble with the lowest R<sub>free</sub>

## METHODS

### B-factor and B'-factor analysis

Isotropic alpha-carbon B-factors were extracted from each structure. For residues with multiple conformers modelled, conformer B was omitted from the analysis. Baseline-normalization was performed using a modified z-transformation method to obtain B'-factors for each alpha carbon ( $B'(i)$ ): Regions of high variance (e.g. loops and termini: residues 0-150, 170-178, 20-210, 230-239, 255-280, 300-450) were omitted when calculating the mean baseline B-factor ( $\bar{B}_{baseline}$ ) and standard deviation of the B-factor ( $\sigma_{baseline}$ ) of each chain prior to calculating  $B'(i)$  (equation 1). Scripts for analysis and plot preparation, as well as full-resolution figures, are available on GitHub.

$$B'(i) = \frac{B(i) - \bar{B}_{baseline}}{\sigma_{baseline}} \quad (1)$$

### Ensemble Refinement

Phenix.ensemble\_refinement<sup>1</sup> (PHENIX version 1.18.2) was used to generate time-averaged ensembles of S1, S2, S3, S5, S6, S7 and S8 (structures from this paper), and R0, R1, R2, R6, R8, R18 and R22 (from previous work<sup>2</sup>). Structures and structure factors were obtained from the wwPDB and re-refined using phenix.refine. For each structure in both Ser and Arg trajectories, phenix.readyset was used to add hydrogen atoms and generate .cif files for the ligands. Each monomer present in the crystallographic dimer was assigned to a unique translation-liberation-screw (TLS) group. Harmonic restraints were added to any ligands present. The ensemble refinement parameters were optimized by testing various TLS fitting procedures ( $p_{TLS} = 0.6, 0.8, 0.9, 1.0$ ) followed by optimization of the X-ray weight (by changing the simulation temperature bath,  $T_{bath} = 290\text{ K}, 295\text{ K}, \text{ and } 297.5\text{ K}$ ) and the relaxation time of the time-averaged restrained ( $\tau_x = 0.3, 0.6, 1$ ). Finally, nine additional random seed repeats were performed using the parameters that led to the ensemble with the lowest  $R_{free}$ . Alpha-carbon RMSF values were calculated using the get\_rmsf tool in GROMACS (version 2018.3); analysis was performed on the five ensembles with the lowest  $R_{free}$  values, following averaging of values across the two chains in the crystallographic dimer. Ensembles were visualized in PyMOL version 2.0 (<http://pymol.org>). Scripts used for analysis, plot preparation and PyMOL figures, as well as full resolution images, are available on GitHub.

## REFERENCES

1. Burnley, B. T., Afonine, P. V., Adams, P. D. & Gros, P. Modelling dynamics in protein crystal structures by ensemble refinement. *Elife* 1, e00311 (2012).
2. Campbell, E. et al. The role of protein dynamics in the evolution of new enzyme function. *Nature Chemical Biology* vol. 12 944–950 (2016).