Biophysical Journal, Volume *115*

Supplemental Information

Mapping the Ligand Binding Landscape

Alex Dickson

Mapping the ligand binding landscape
Supplemental Information

Alex Dickson*
Department of Biochemistry & Molecular Biology, Michigan State University, East Lansing, MI 48824 and Department of Computational Mathematics, Science and Engineering, Michigan State University, East Lansing, MI 48824
(Dated: September 26, 2018)

I. SUPPLEMENTAL FIGURES

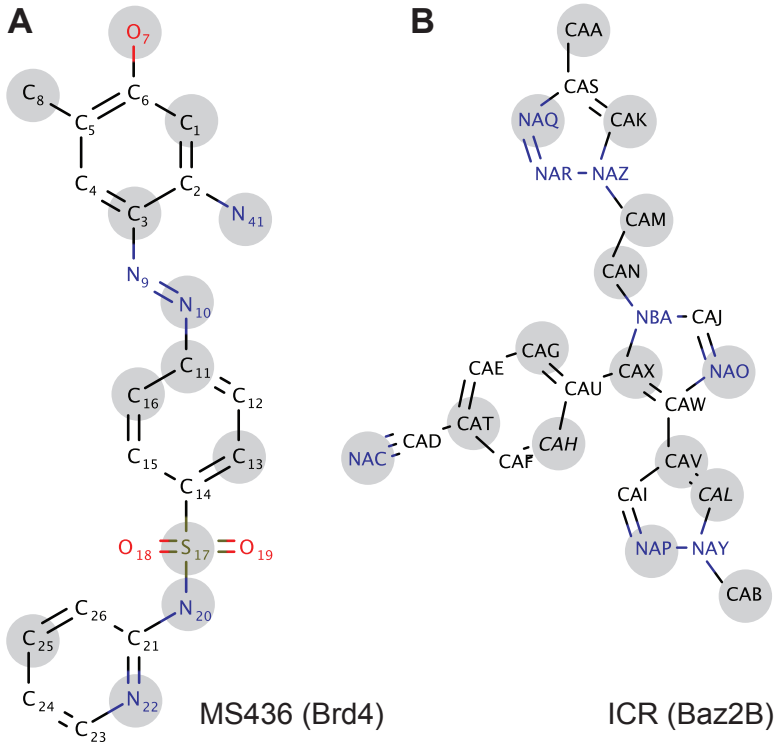

FIG. S1: **Structure of ligands**. Chemical structure is shown for both ligands examined in this study. Atoms are labeled according to their name in the PDB. Atoms highlighted with grey circles are selected to construct the set of ligand-protein distances used for conformation space network analysis. **A)** The MS436 ligand of Brd4, labeled as in PDB structure 4NUD [1]. **B)** The ICR ligand of Baz2B, labeled as in PDB structure 4XUB [2]. Both ligand structures were prepared using MarvinSketch.

*Electronic address: alexrd@msu.edu

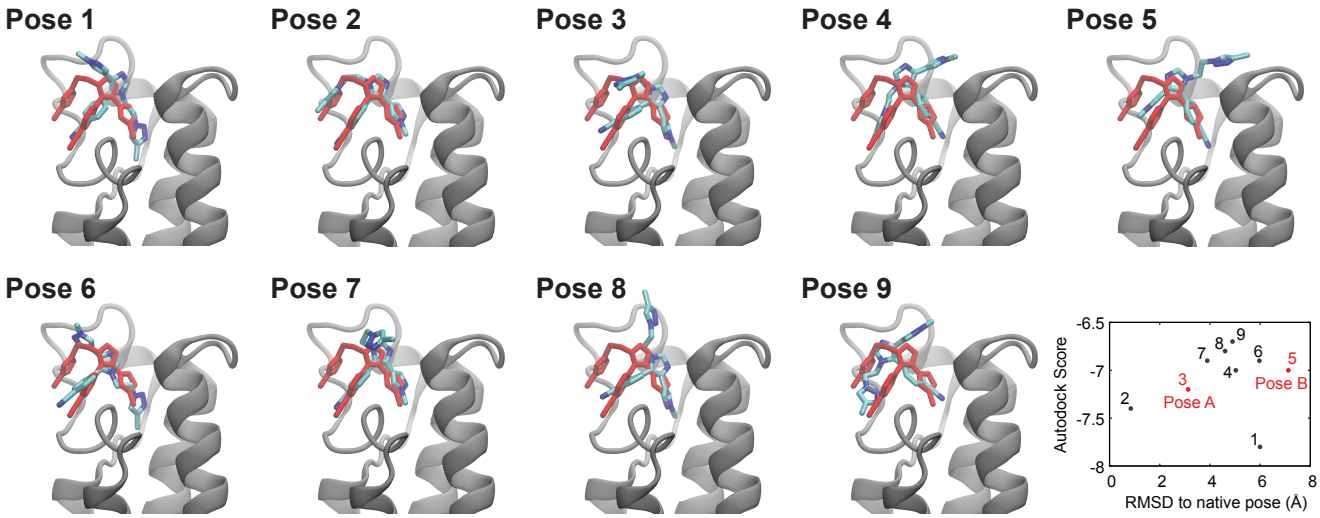

FIG. S2: **Top nine poses from docking for Baz2B-ICR**. All nine poses are colored by atom type and overlayed onto the crystal structure, shown in red. In the bottom right panel, a scatter plot shows the relationship between the free energy score from Autodock Vina [3] and the root mean squared displacement (RMSD) from the native pose, computed using all atoms except hydrogen. Pose 1 shows the best score, but is 6 Å from the native pose. Pose 2 is within 1 Å of the crystal structure. For starting points in our WExplore simulations, we choose Pose 3 (referred to in the text as Pose A) and Pose 5 (referred to in the text as Pose B).

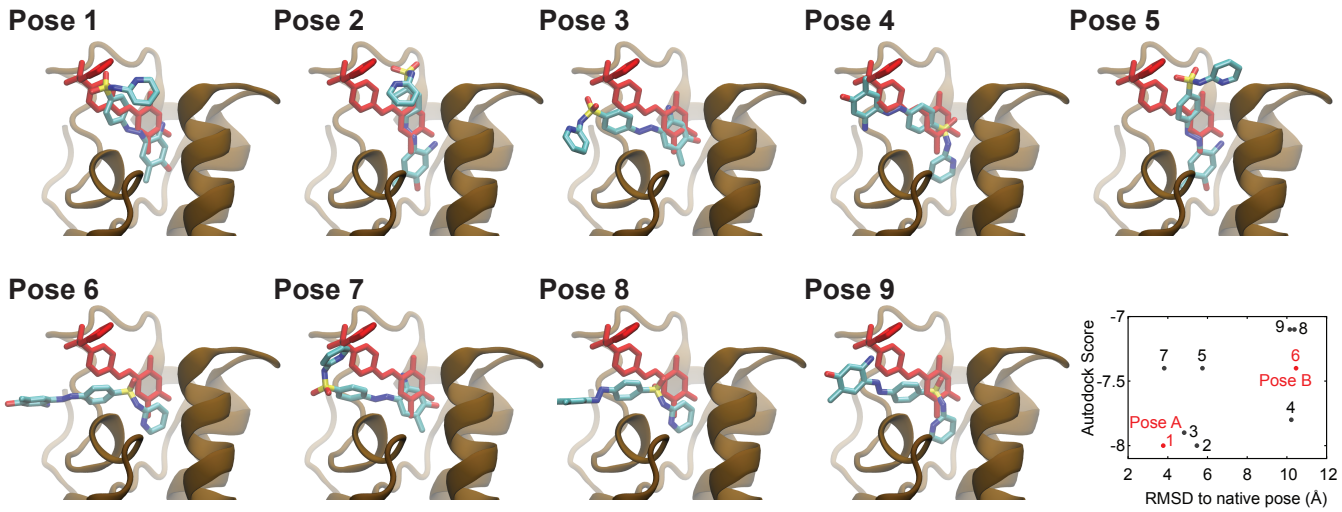

FIG. S3: **Top nine poses from docking for Brd4-MS436**. All nine poses are colored by atom type and overlayed onto the crystal structure, shown in red. The bottom right panel shows a scatter plot of the free energy score from Autodock Vina [3] and the native RMSD, again computed using all atoms except hydrogen. None of the nine poses showed a good agreement to the crystal structure pose. Poses 1 and 7 are the closest, both with RMSD = 3.8 Å. For starting points in our WExplore simulations, we choose Pose 1 (referred to in the text as Pose A) and Pose 6 (referred to in the text as Pose B).

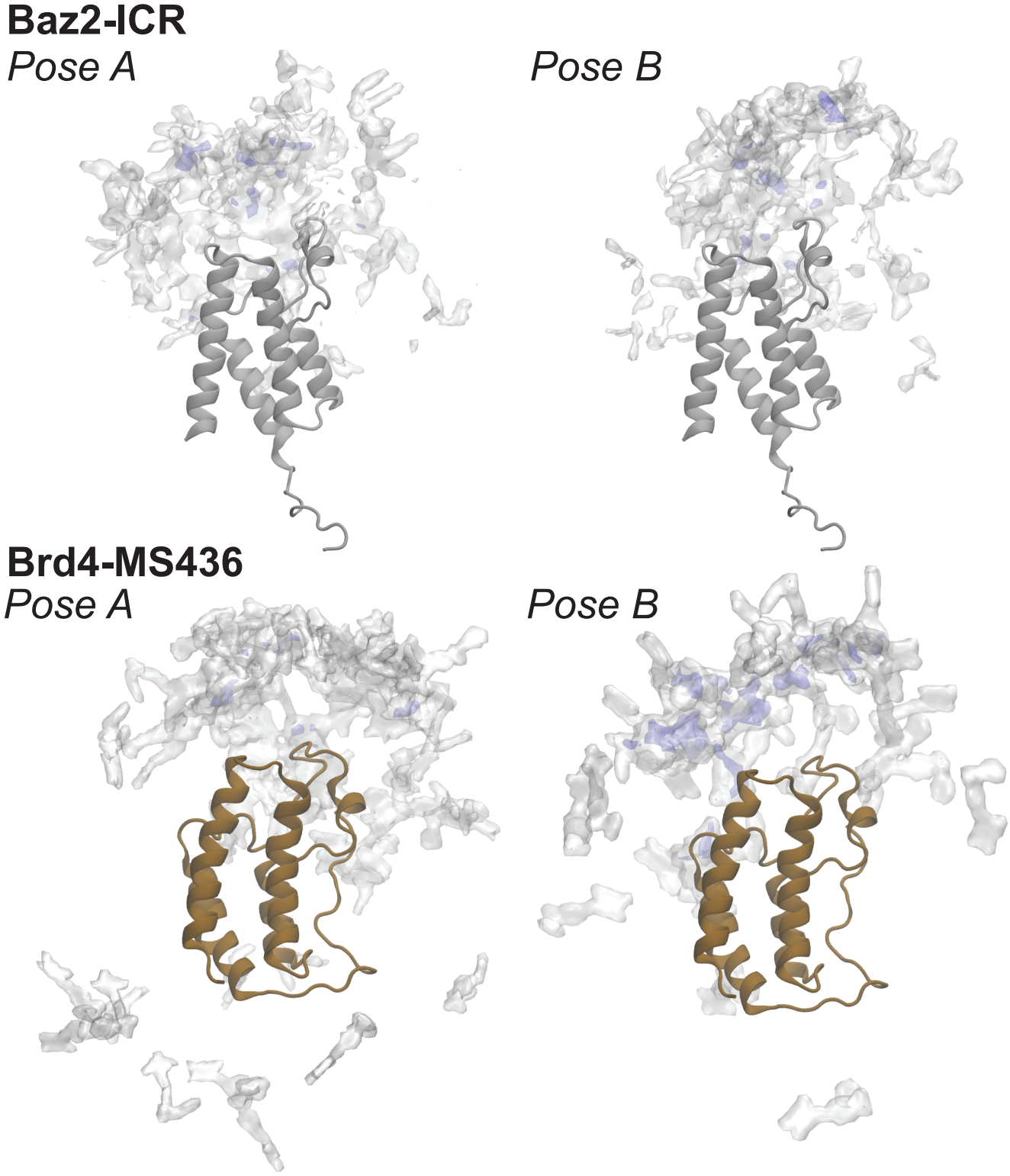

FIG. S4: **Exit point clouds**. For Baz2B-ICR7 (top) and Brd4-MS436 (bottom), we show density distributions for ligand exit points, which are conformations where the closest interatomic distance between the ligand and protein is at least 10 Å. The densities are determined using the VolMap facility in VMD, and are visualized using isovalues 0.02 (transparent white) and 0.08 (transparent blue).

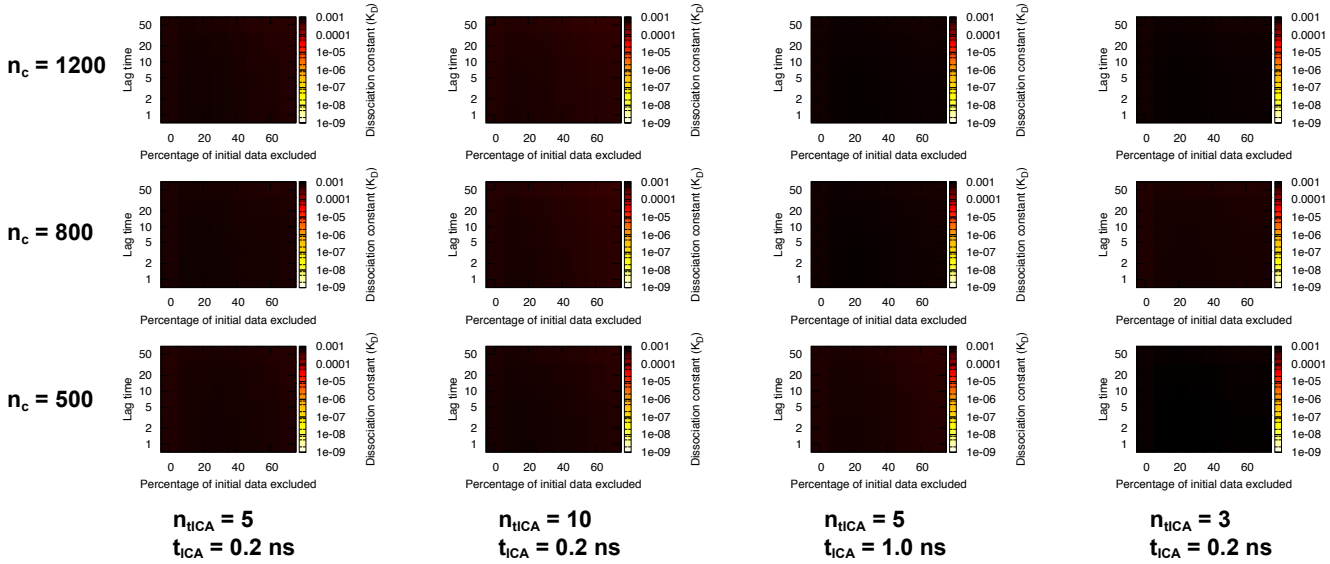

FIG. S5: **Dissociation constants for Baz2B calculated with unweighted transition counts**. Markov state model parameters used are given in the row and column labels, and each grid shows dissociation constants calculated with 8 different p_{excl} values (0, 10, 20, 30, 40, 50, 60, 70) and 6 different Markov model lag times (1, 2, 5, 10, 20, 50) in units of Δt.

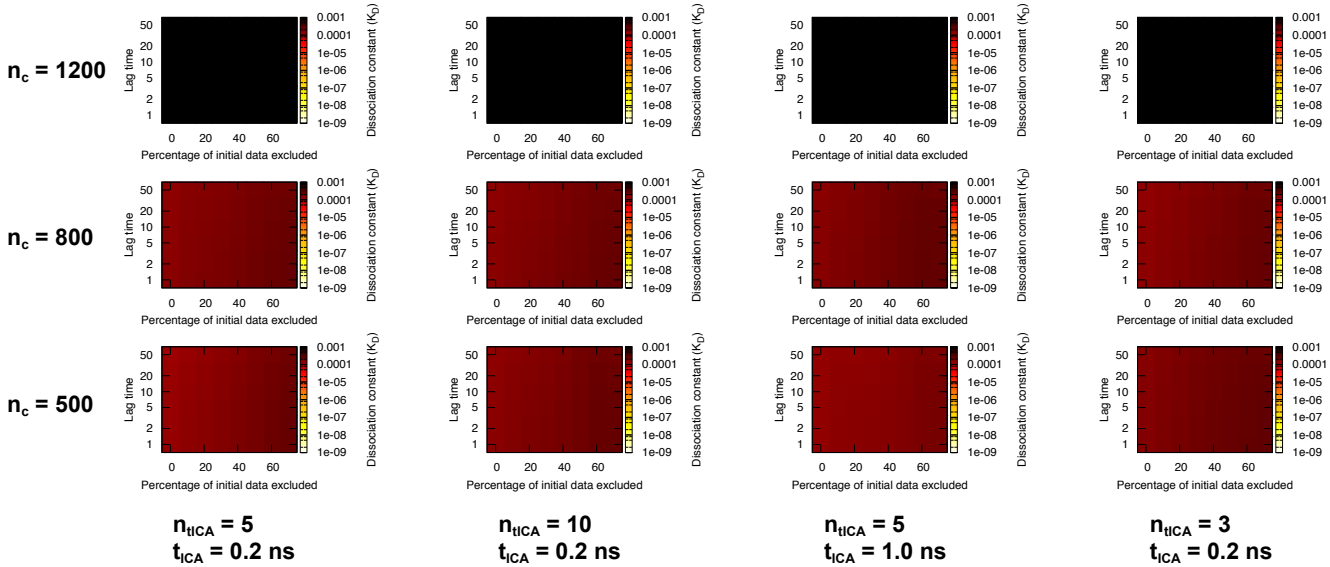

FIG. S6: **Dissociation constants for Brd4 calculated with unweighted transition counts**. Markov state model parameters used are given in the row and column labels, and each grid shows dissociation constants calculated with 8 different p_{excl} values (0, 10, 20, 30, 40, 50, 60, 70) and 6 different Markov model lag times (1, 2, 5, 10, 20, 50) in units of Δt.

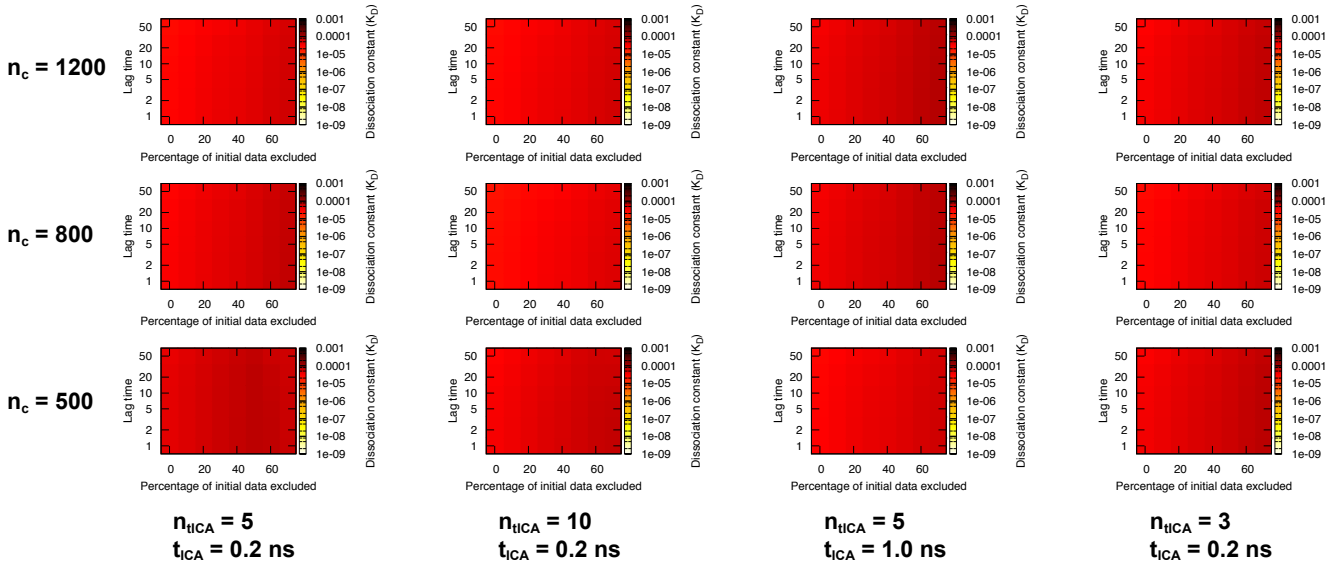

FIG. S7: **Dissociation constants for Baz2B calculated with weighted transition counts**. Markov state model parameters used are given in the row and column labels, and each grid shows dissociation constants calculated with 8 different p_{excl} values (0, 10, 20, 30, 40, 50, 60, 70) and 6 different Markov model lag times (1, 2, 5, 10, 20, 50) in units of Δt.

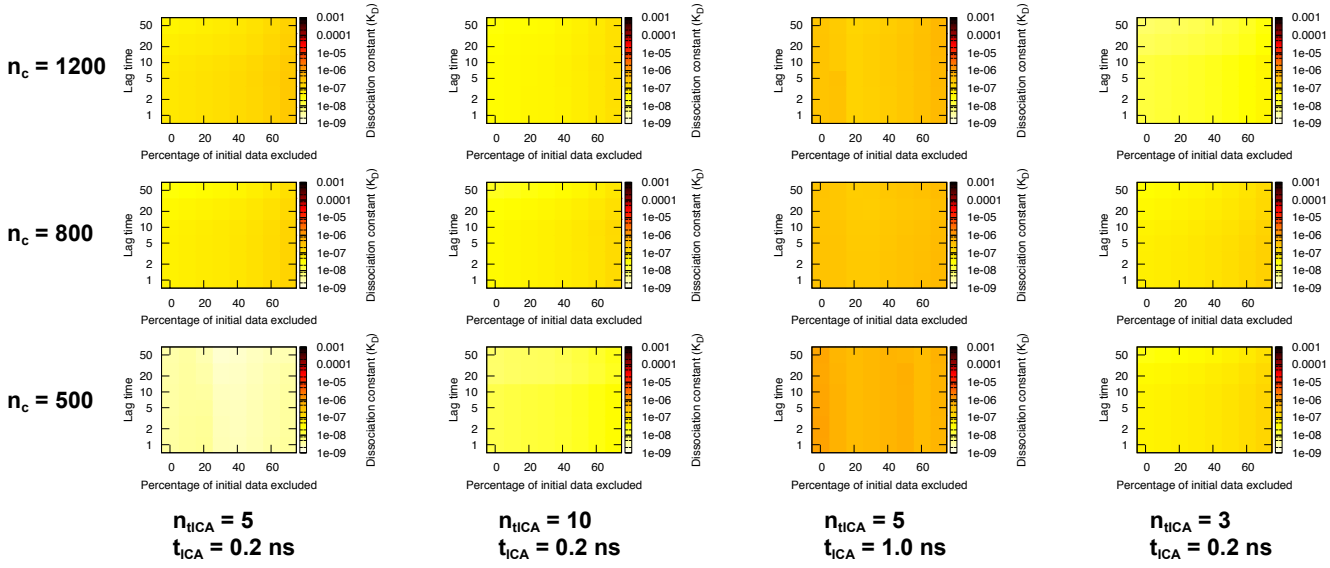

FIG. S8: **Dissociation constants for Brd4 calculated with weighted transition counts**. Markov state model parameters used are given in the row and column labels, and each grid shows dissociation constants calculated with 8 different p_{excl} values (0, 10, 20, 30, 40, 50, 60, 70) and 6 different Markov model lag times (1, 2, 5, 10, 20, 50) in units of Δt.

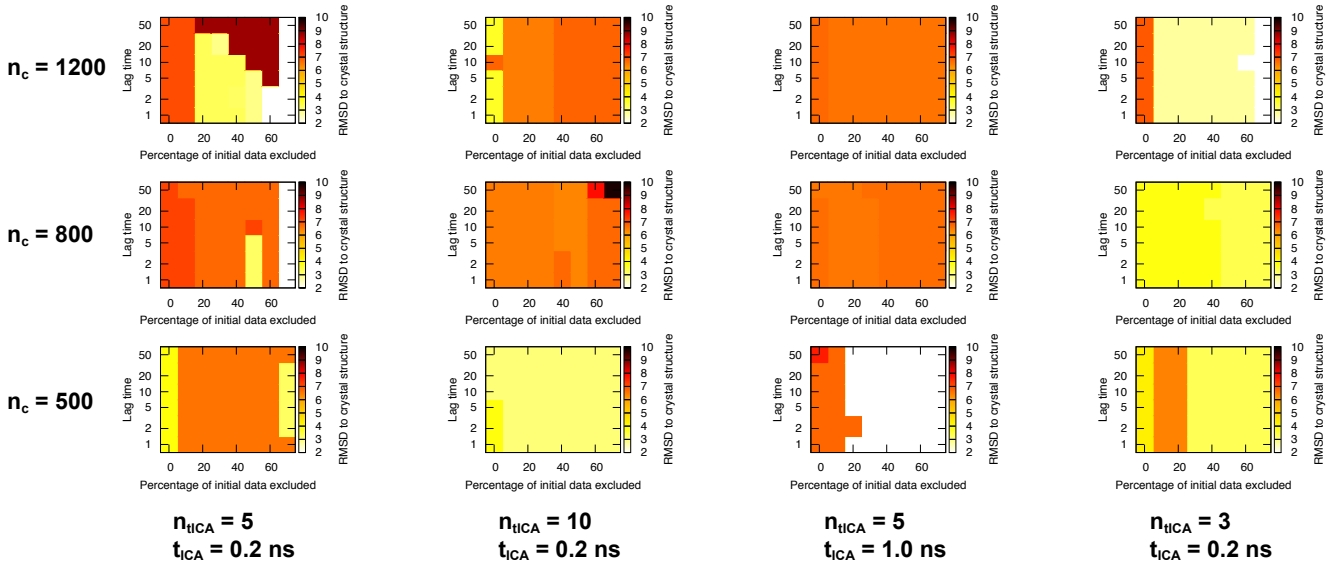

FIG. S9: **RMSD of lowest free-energy pose for Baz2B calculated with unweighted transition counts**. Markov state model parameters used are given in the row and column labels, and each grid shows dissociation constants calculated with 8 different p_{excl} values (0, 10, 20, 30, 40, 50, 60, 70) and 6 different Markov model lag times (1, 2, 5, 10, 20, 50) in units of Δt.

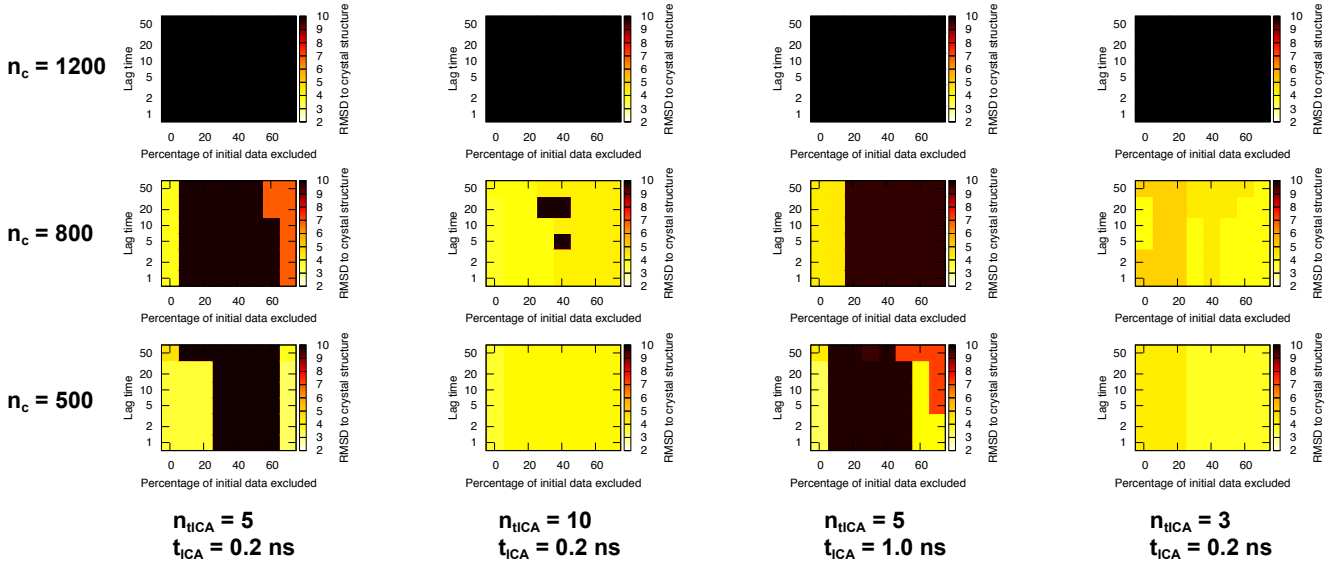

FIG. S10: **RMSD of lowest free-energy pose for Brd4 calculated with unweighted transition counts**. Markov state model parameters used are given in the row and column labels, and each grid shows dissociation constants calculated with 8 different p_{excl} values (0, 10, 20, 30, 40, 50, 60, 70) and 6 different Markov model lag times (1, 2, 5, 10, 20, 50) in units of Δt.

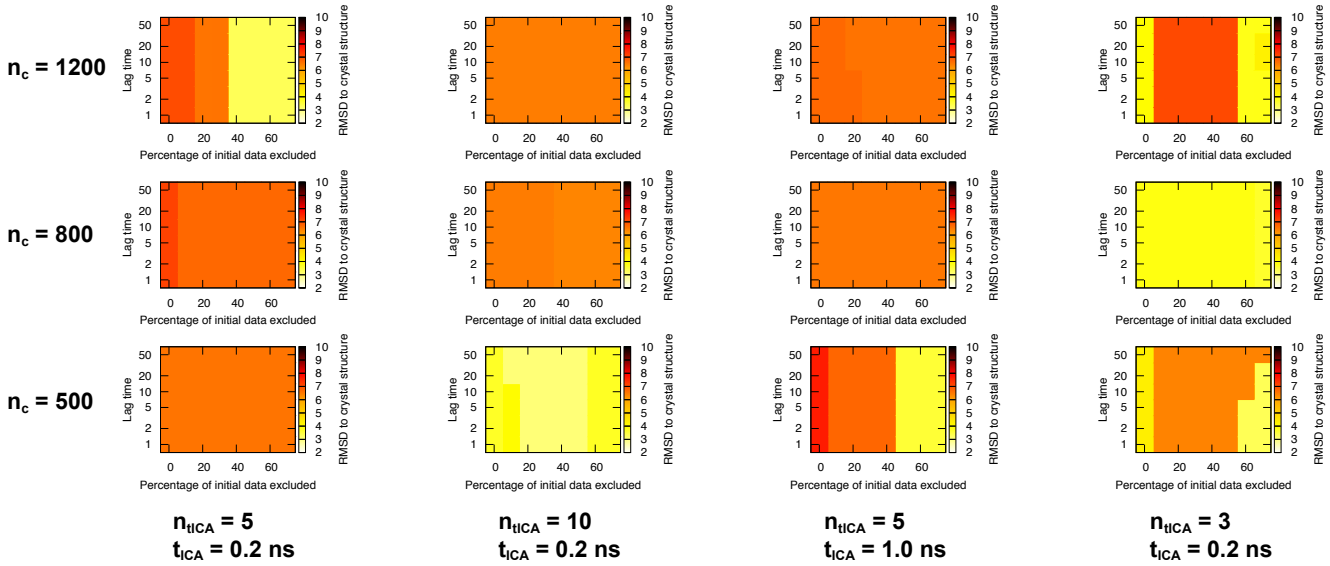

FIG. S11: **RMSD of lowest free-energy pose for Baz2B calculated with weighted transition counts**. Markov state model parameters used are given in the row and column labels, and each grid shows dissociation constants calculated with 8 different p_{excl} values (0, 10, 20, 30, 40, 50, 60, 70) and 6 different Markov model lag times (1, 2, 5, 10, 20, 50) in units of Δt.

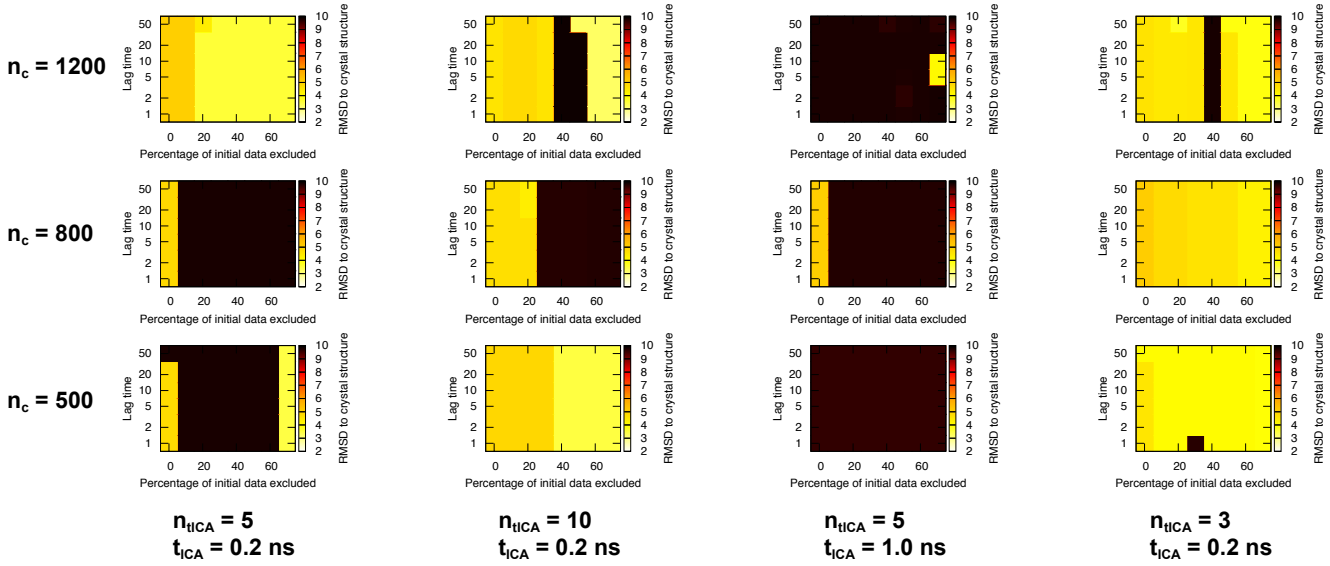

FIG. S12: **RMSD of lowest free-energy pose for Brd4 calculated with weighted transition counts**. Markov state model parameters used are given in the row and column labels, and each grid shows dissociation constants calculated with 8 different p_{excl} values (0, 10, 20, 30, 40, 50, 60, 70) and 6 different Markov model lag times (1, 2, 5, 10, 20, 50) in units of Δt.

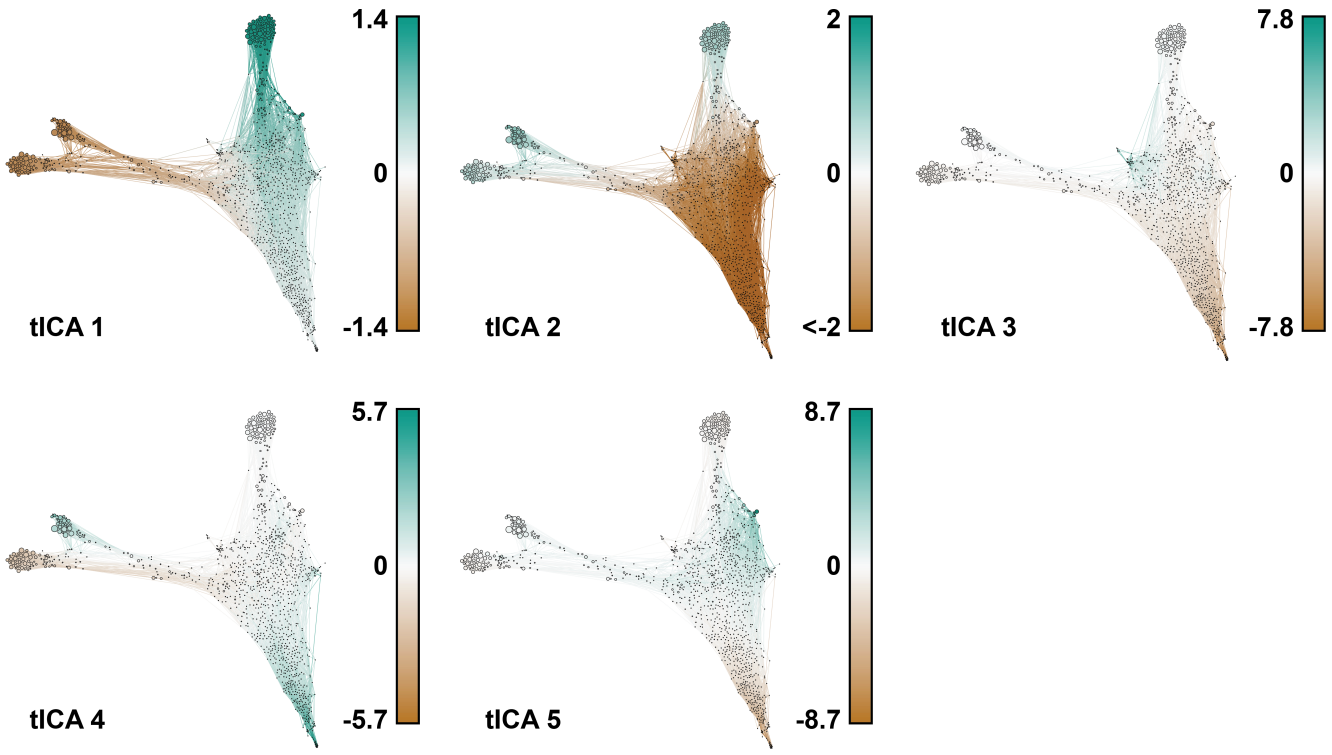

FIG. S13: **Collective variables from time-lagged independent component analysis for Baz2B-ICR7**. Five tICA variables are used to color the pose networks for the Baz2B-ICR7 network. The color scale is symmetric in each case, but was clipped for tICA 2 to better show the states with positive tICA values (green).

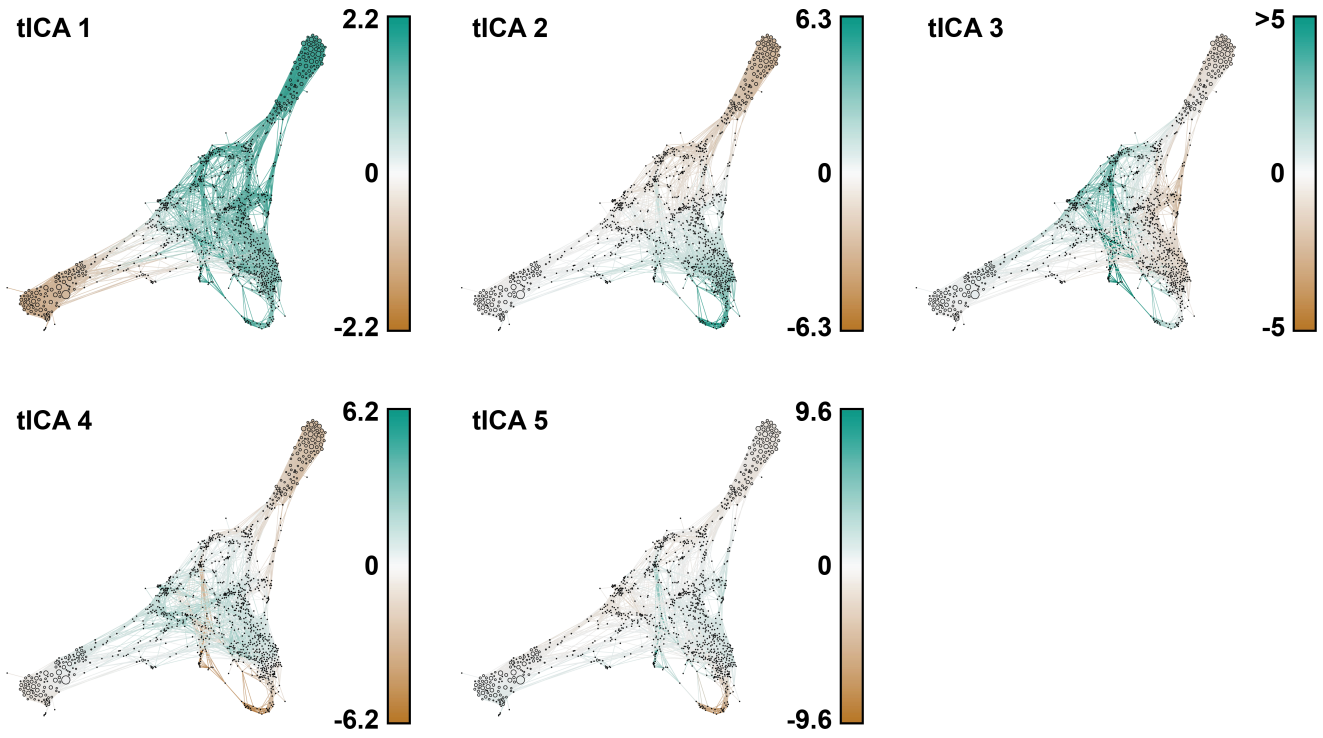

FIG. S14: **Collective variables from time-lagged independent component analysis for Brd4-MS436**. Five tICA variables are used to color the pose networks for the Brd4-MS436 network. The color scale is symmetric in each case, but was clipped for tICA 3 to better show the states with negative tICA values (brown).

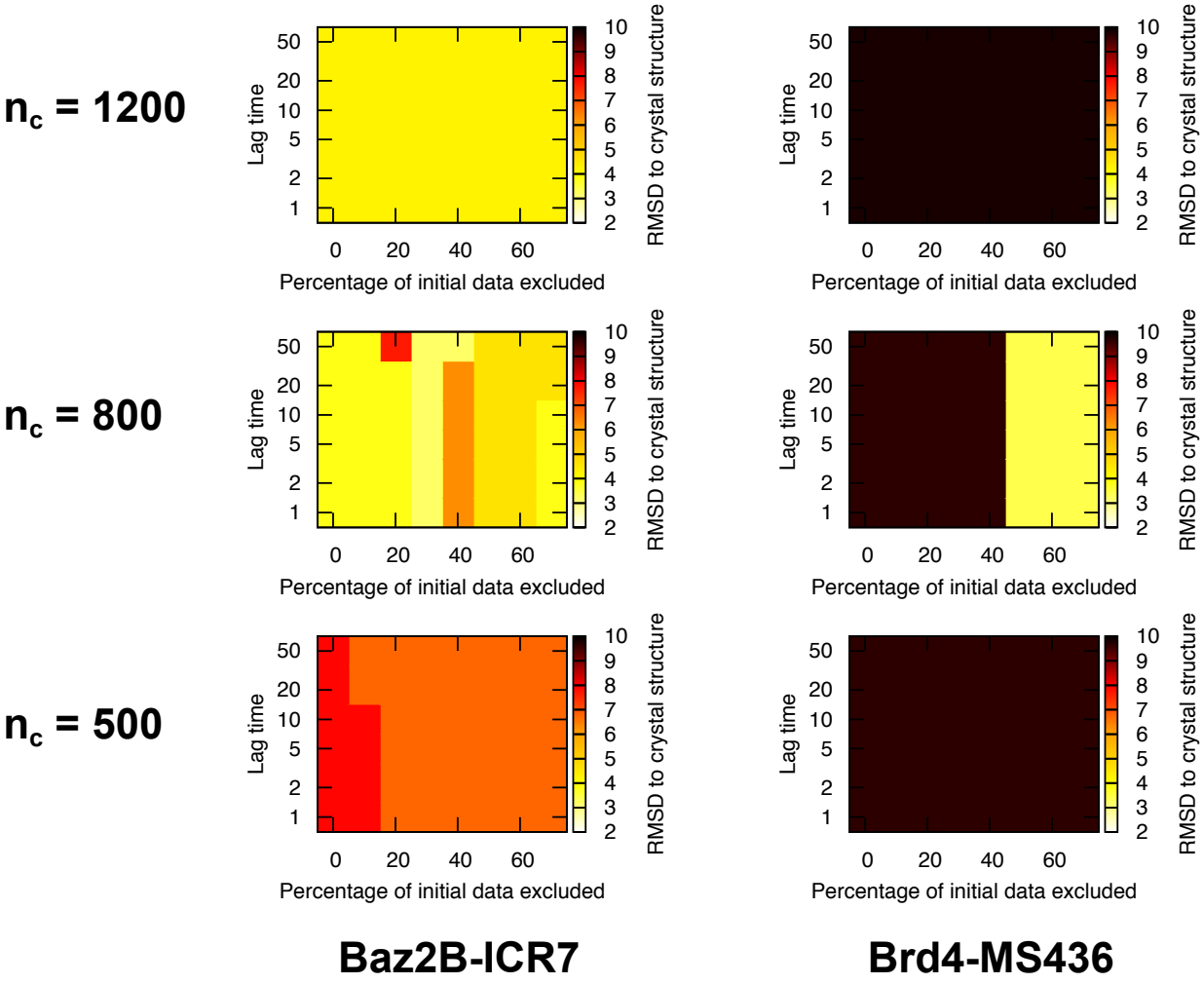

FIG. S15: **RMSD of lowest free energy cluster without tICA**. Average RMSD values are calculated using 10 randomly-chosen structures from the state with the lowest free energy for each set of Markov model parameters. Each heat map shows 48 results from 48 different Markov state models.

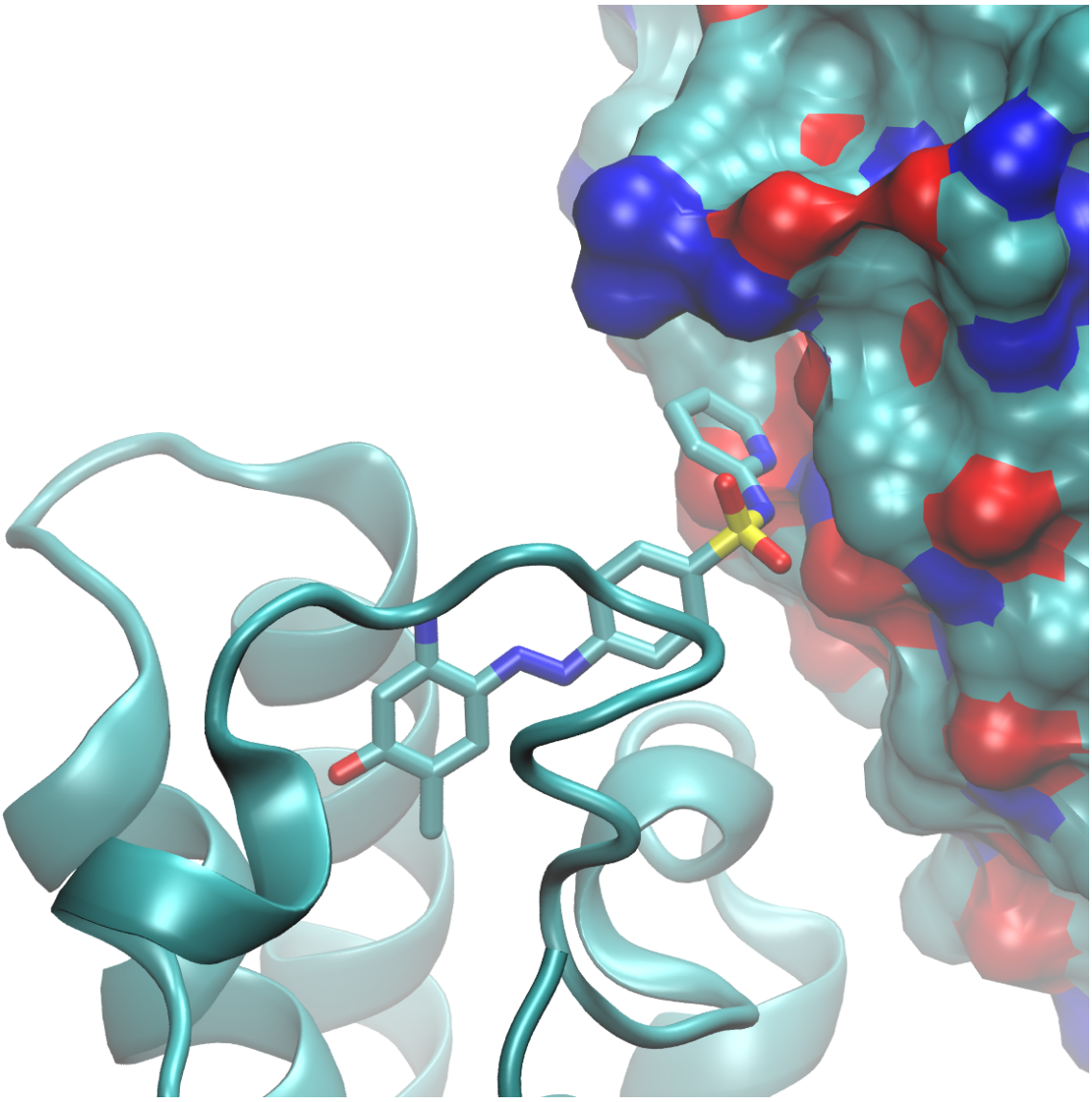

FIG. S16: **Ligand associates with multiple copies of Brd4 in PDB 4NUD**. One copy of Brd4 is shown in cartoon representation, and the other is shown in surface representation.

II. SUPPLEMENTAL TABLES

[1] G. Zhang, A. N. Plotnikov, E. Rusinova, T. Shen, K. Morohashi, J. Joshua, L. Zeng, S. Mujtaba, M. Ohlmeyer, and M. M. Zhou, Journal of Medicinal Chemistry **56**, 9251 (2013).
[2] L. Drouin, S. McGrath, L. R. Vidler, A. Chaikuad, O. Monteiro, C. Tallant, M. Philpott, C. Rogers, O. Fedorov, M. Liu, et al., Journal of Medicinal Chemistry **58**, 2553 (2015).
[3] O. Trott and A. Olson, Journal of Computational Chemistry **31**, 455 (2010).

TABLE S1: Parameter sets used for clustering with tICA.

n_{tICA}	τ_{tICA} (ns)	n_c
3	0.2	500
		800
		1200
5	0.2	500
		800
		1200
5	1.0	500
		800
		1200
10	0.2	500
		800
		1200

TABLE S2: Calculated diffusion constants (K_D)

		average (μM)	min(μM)	max (μM)
Baz2B-ICR7	unweighted	740	500	1000
	weighted	47	22	86
	IC50 [a]	1.1		
Brd4-MS436	unweighted	190	120	270
	weighted	0.042	0.0013	0.84
	K_i [b]	0.03-0.05		

[a] Measured in Ref. [2]
[b] Measured in Ref. [1]