Biophysical Journal, Volume 97

Supporting Material

Collective Dynamics of Periplasmic Glutamine Binding Protein upon Domain Closure

Hannes H. Loeffler and Akio Kitao

Collective Dynamics of Periplasmic Glutamine Binding Protein upon Domain Closure (Supplement)

Hannes H. Loeffler*,[1] and Akio Kitao[1,2]
[1]Laboratory of Molecular Design,
Institute of Molecular and Cellular Biosciences,
University of Tokyo
[2]JST, CREST

August 12, 2009

1 Overview

This Supplement discusses supporting information relating to material in the main text. In the first part more detailed insights into the results presented in the main text are discussed. The second part lists additional tables and figures.

2 Results and Discussion

2.1 General Structure

To assess basic dynamics and stability of the simulations we analyze the root mean square deviations (RMSD) in the following. Fig. S1 shows the RMSD for simulations of the four states in two box sizes. As reference structures the closed crystal structure 1WDN for the closed states and the open crystal structure 1GGG for the open states were chosen.

The closed-liganded state CL is the one with the least structural fluctuations and therefore the most stable with respect to the crystal structure. The open-liganded state OL shows a larger RMSD and fluctuations due to larger domain motions (see below). Removing the ligand leads to a less stable structure in the close-unliganded case CU as evidenced from its RMSD which is quite similar to OL but the fluctuations are smaller than OL. The most flexible structure appears in the open-unliganded simulation OU. Based on the last 25 ns of simulation and a sampling rate of 1 ps the RMSD statistics is summarized in Table S2.

The larger box sizes with 140 000 atoms display RMSD fluctuations that are quite similar to the smaller box sizes. The CL simulations are closest to each other with virtually the same statistics. The larger box OL simulation

*Corresponding author: hannes.loeffler@stfc.ac.uk (H. H. Loeffler). Current address: STFC Daresbury Laboratory, Warrington WA4 4AD, United Kingdom.

appears to be somewhat more dynamic than the smaller box but statistics is quite similar.

Analysis of the RMSD of individual domains (see Fig. S2), i.e. of the small domain and the large domain, show that the large domain typically deviates less from the respective crystal structure than the small domain. The exceptions are the case of the two closed-liganded simulations where the two domain RMSDs and therefore also the total RMSD are very similar to each other. In the OU and big box OL simulations the large domain RMSD is initially larger but becomes similar to the small domain RMSD later in the simulation. Fluctuations in the domain RMSDs are rather small. They are notable smaller than the total RMSD in the open states. This means that the two domains may be considered rigid bodies with respect to global movements of the protein around the hinge.

The root mean square fluctuations (RMSF) computed per residue and averaged over time are depicted in Fig. S3. The general picture of the smaller box sizes is that regions of strong and weak RMSFs are similar in all four states. The largest RMSFs appear in both the N- and the C-terminus but also in a loop region around residue 20 which was suspected to be involved in GlnBP–GlnP receptor binding (1). Other regions of large RMSF are a loop between approximately residues 100 to 110, and the region between residues 115 and 150. Both regions are located in the small domain. This is consistent with the picture drawn by the RMSDs of individual domains, i.e. the RMSFs pinpoint the location of largest fluctuations in the small domain.

Relatively small RMSFs can be found in the hinge regions residues 85–89 and residues 181–185 with the exception of the OU simulation. The closed states exhibit smaller values than the open states consistent with larger domain motions in the open states. The smallest RMSFs, however, are located in residues number just above the hinge regions (residues 92–93 and 187). RMSFs of residues binding to the ligand show a somewhat mixed behavior. Residues Asp10 and Phe13 have values smaller than 1.2 Å in all states whereas Phe50 is between 1.0 Å and 1.7 Å. Residues Ala67, Gly68, Ile69 and Thr70 are smaller than 0.9 Å in the CL, CU and OL cases. Arg75, a key residue in ligand binding (1), is small (0.85 and 0.9 Å, respectively) in the liganded states but larger in the unliganded states (1.0 and 1.2 Å). Lys115 and Gly119 of the small domain fluctuate little in the closed-liganded state (0.6 and 0.7 Å, respectively) but much more pronounced in the other states (1.0–1.6 Å). His156 and Asp157 exhibit RMSF values smaller than 1.0 Å in all cases.

The main difference in RMSF between the smaller and larger box sizes (see Fig. S3) in the two closed-liganded state simulations is the larger fluctuation of the loop around residue 105. The overall picture of the two open-liganded cases is that the larger OL system allows somewhat larger fluctuations in some regions whereas the two CL simulations are nearly identical to each other.

2.2 Collective Motions

To assess the convergence of the eigenvalues we split the last 25 ns of the trajectories into five 5 ns patches and recalculated the PCAs (Fig. S5). The spread in the distribution of eigenvalues is relatively small in the cases of the CL, CU and OL simulations but larger in the three other simulations. A clear trend or picture cannot be drawn but it is obvious that the eigenvalues between adjacent 5 ns patches may be significantly altered by a factor of up to two.

However, sufficient convergence of the eigenvalues alone does not necessarily mean that the eigenvectors have converged and vice versa as has been pointed out earlier (2,3). In principal, one can try to assess the convergence of a subspace by calculating the overlap function (2).

In our case the subspace spanned by the first 25 eigenvectors converges within a few nanoseconds (see Fig. S6). However, we are more interested in the similarity of individual modes, i.e. for instance if the first modes of the open state simulations indicate similar global motions. To do so we computed the dot products for single modes. A value close to 1 means that both eigenvectors point in the same direction, a value close to 0 implies orthogonality and thus low similarity.

It is obvious from Table S2 that the highest similarity with 0.81 can be found between the first modes of the two open state simulations. Also the second modes show some similarity of 0.65 but in all other cases eigenvectors describe different motions. Comparing the larger boxes with the smaller box simulations we find 0.96 for the OL case and 0.55 for the CL case, (both numbers for 1^{st} mode).

Similar to the comparisons above we can also compare the eigenvector dot products between 5 ns patches and modes. Tables S4 to S9 summarize dot products of the same mode. The CL and the CU simulations typically show low similarity between patches but some dot products highlight similar modes although not necessarily between directly adjacent patches. The OL and the OU simulations are highly similar between nearly all patches in both 1^{st} and 2^{nd} modes. Comparison with the corresponding eigenvalues (see Fig. S5) shows that different magnitudes on the one hand and different dot products on the other hand do not appear to correlate.

Cross-dot products (see Tables S10 to S15) show some higher values between 1^{st} and 2^{nd} modes for CL and CU simulations indicating some degree of similarity but a clear mode reversal can not be observed. The OU and OL systems on the other hand exhibits very low similarity close to orthogonality between 1^{st} and 2^{nd} modes. The bigger box sizes behave like their smaller box counterparts.

The projections of the first three, the 10^{th} and the 25^{th} principal component vectors on to the trajectories together with the probability distributions p and the free energy $G = -kT \ln p$ (in kJ/mol) along the principal mode are shown in the Figs. S9 to S8. Despite the largest amplitude modes not being harmonic (4) we find relatively Gaussian like distributions in most projections indicative of well converged results. The larger box sizes exhibit less well defined distributions in the slowest modes.

2.3 Secondary Structure and Hydrogen Bonds

The secondary structure is mostly stable in all simulations. Typical exceptions to this observation are the terminal regions of secondary structure elements which show a larger flexibility than the rigid core residues. Larger variability in secondary structure can be found for the loop terminated by residues 96 and 110. This loop partially forms a 3-10 helix in some simulations. But this helix is not stable and only of transient nature.

Various hydrogen bonds of the open-unliganded (5) and closed-liganded (1)

conformations are discussed in the original publications on the crystallographic structure resolution. The closed-liganded crystal structure shows 13 hydrogens bonds in total between the glutamine ligand and its neighbors, 9 to the large domain and 4 to the small domain. In the corresponding MD simulations we find the same number of H-bonds minus the one between Gln227:OE1 and His156:NE2 (small domain). Also, donors and acceptors are different in some cases. Thr70:OG1, for instance, hydrogen-binds to Gln183:OE1 rather than to Gln183:NE2 as predicted from close-liganded 1WDN, i.e. the amid group of Gln183 flips over during the simulation. In contrast, both chains of 1GGG have the amid group flipped with respect to 1WDN.

The glutamine ligand does not directly contact the hinges through hydrogen bonding but rather two hydrogen bond networks are formed to the same hinge strand. One is as described above, i.e. α-carboxy O Gln227 to Thr70:OG1 to Gln183:OE1 according to the MD simulations. The other one is α-amino N Gln227 to Asp157:OD to Tyr185:OH where Asp157 from the small domain will naturally only bind in the closed state.

References

1. Sun, Y.-J., J. Rose, B.-C. Wang, and C.-D. Hsiao, 1998. The Structure of Glutamin-binding Protein Complexed with Glutamine at 1.94 Å Resolution: Comparisons with other Amino Acid Binding Proteins. *J. Mol. Biol.* 279:219–229.

2. Hess, B., 2000. Similarities between principal components of protein dynamics and random diffusion. *Phys. Rev. E* 62:8438–8448.

3. Hess, B., 2002. Convergence of sampling in protein simulations. *Phys. Rev. E* 65:031910.

4. Lars Meinhold, J. C. S., 2007. Protein dynamics from X-ray crystallography: Anisotropic, global motion in diffuse scattering patterns. *Proteins: Struct. Func. Bioinf.* 66:941–953.

5. Hsiao, C.-D., Y.-J. Sun, J. Rose, and B.-C. Wang, 1996. The Crystal Structure of Glutamine-binding Protein from Escherichia coli. *J. Mol. Biol.* 262:225–242.

3 Tables

Table S1: Statistics from the RMSD as calculated from the last 25 ns and a sampling rate of 1 ps. All values in Å.

simulation	minimum	maximum	average	σ^a
CL	0.99	2.19	1.37	0.16
CU	1.31	2.67	1.82	0.18
OL	1.19	2.79	1.76	0.25
OU	1.23	2.91	1.93	0.35
CL big	0.94	2.14	1.44	0.16
OL big	1.04	3.02	1.73	0.29

[a]standard deviation

Table S2: Dot products of individual modes indicating similarity for the first two modes and the smaller box sizes. Lower triangle of the matrix is the same as the upper triangle. Diagonal elements are all 1.

	1^{st} mode				2^{nd} mode			
	CL	CU	OL	OU	CL	CU	OL	OU
CL		0.40	0.31	0.17		0.51	0.15	0.05
CU			0.15	0.05			0.52	0.08
OL				0.47				0.13

Table S3: Eigenmode cross correlations between the first two modes.

	$1^{st}/2^{nd}$		$2^{nd}/1^{st}$	
	OU	OL	OU	OL
CU	0.15	0.18	0.54	0.53
CL	0.39	0.58	0.02	0.42

Table S4: Eigenmode dot products of 5 ns patches for the CL simulation. Lower triangle of the matrix is the same as the upper triangle. Diagonal elements are all 1.

	1^{st} mode					2^{nd} mode				
	1	2	3	4	5	1	2	3	4	5
6–10 ns, 1		0.31	0.60	0.10	0.50		0.35	0.50	0.03	0.51
11–15 ns, 2			0.71	0.23	0.46			0.85	0.26	0.80
16–20 ns, 3				0.23	0.73				0.19	0.77
21–25 ns, 4					0.43					0.51

Table S5: Eigenmode dot products of 5 ns patches for the CU simulation. Lower triangle of the matrix is the same as the upper triangle. Diagonal elements are all 1.

	1^{st} mode					2^{nd} mode				
	1	2	3	4	5	1	2	3	4	5
6–10 ns, 1		0.22	0.46	0.49	0.51		0.18	0.57	0.01	0.11
11–15 ns, 2			0.75	0.68	0.05			0.01	0.55	0.61
16–20 ns, 3				0.79	0.08				0.28	0.08
21–25 ns, 4					0.15					0.34

Table S6: Eigenmode dot products of 5 ns patches for the OL simulation. Lower triangle of the matrix is the same as the upper triangle. Diagonal elements are all 1.

	1^{st} mode					2^{nd} mode				
	1	2	3	4	5	1	2	3	4	5
6–10 ns, 1		0.95	0.90	0.91	0.95		0.80	0.76	0.75	0.80
11–15 ns, 2			0.93	0.89	0.96			0.91	0.85	0.89
16–20 ns, 3				0.86	0.92				0.83	0.83
21–25 ns, 4					0.91					0.90

Table S7: Eigenmode dot products of 5 ns patches for the OU simulation. Lower triangle of the matrix is the same as the upper triangle. Diagonal elements are all 1.

	1^{st} mode					2^{nd} mode				
	1	2	3	4	5	1	2	3	4	5
6–10 ns, 1		0.92	0.89	0.92	0.81		0.72	0.66	0.47	0.07
11–15 ns, 2			0.83	0.96	0.81			0.80	0.76	0.04
16–20 ns, 3				0.83	0.67				0.61	0.16
21–25 ns, 4					0.88					0.08

Table S8: Eigenmode dot products of 5 ns patches for the CL big simulation. Lower triangle of the matrix is the same as the upper triangle. Diagonal elements are all 1.

	1^{st} mode					2^{nd} mode				
	1	2	3	4	5	1	2	3	4	5
6–10 ns, 1		0.84	0.21	0.44	0.54		0.18	0.40	0.78	0.82
11–15 ns, 2			0.40	0.49	0.56			0.30	0.05	0.04
16–20 ns, 3				0.34	0.28				0.39	0.31
21–25 ns, 4					0.49					0.80

Table S9: Eigenmode dot products of 5 ns patches for the OL big simulation. Lower triangle of the matrix is the same as the upper triangle. Diagonal elements are all 1.

	1^{st} mode					2^{nd} mode				
	1	2	3	4	5	1	2	3	4	5
6–10 ns, 1		0.92	0.83	0.89	0.95		0.72	0.63	0.54	0.09
11–15 ns, 2			0.84	0.89	0.87			0.85	0.79	0.49
16–20 ns, 3				0.89	0.86				0.74	0.38
21–25 ns, 4					0.91					0.23

Table S10: Eigenmode cross dot products of 5 ns patches for the CL simulation.

	1^{st} mode/2^{nd} mode				
	1	2	3	4	5
6–10 ns, 1	0.00	0.51	0.43	0.73	0.65
11–15 ns, 2	0.05	0.00	0.00	0.46	0.36
16–20 ns, 3	0.02	0.08	0.00	0.73	0.45
21–25 ns, 4	0.30	0.48	0.66	0.00	0.30
26–30 ns, 5	0.33	0.19	0.31	0.59	0.00

Table S11: Eigenmode cross dot products of 5 ns patches for the CU simulation.

	1^{st} mode/2^{nd} mode				
	1	2	3	4	5
6–10 ns, 1	0.00	0.77	0.18	0.33	0.65
11–15 ns, 2	0.09	0.00	0.29	0.35	0.49
16–20 ns, 3	0.00	0.34	0.00	0.09	0.64
21–25 ns, 4	0.27	0.32	0.45	0.00	0.57
26–30 ns, 5	0.02	0.32	0.28	0.10	0.00

Table S12: Eigenmode cross dot products of 5 ns patches for the OL simulation.

	1^{st} mode/2^{nd} mode				
	1	2	3	4	5
6–10 ns, 1	0.00	0.06	0.32	0.09	0.03
11–15 ns, 2	0.01	0.00	0.25	0.18	0.12
16–20 ns, 3	0.15	0.26	0.00	0.39	0.32
21–25 ns, 4	0.14	0.07	0.34	0.00	0.03
26–30 ns, 5	0.11	0.08	0.33	0.08	0.00

Table S13: Eigenmode cross dot products of 5 ns patches for the OU simulation.

	1^{st} mode/2^{nd} mode				
	1	2	3	4	5
6–10 ns, 1	0.00	0.20	0.19	0.23	0.32
11–15 ns, 2	0.14	0.00	0.42	0.07	0.43
16–20 ns, 3	0.17	0.36	0.00	0.36	0.38
21–25 ns, 4	0.16	0.06	0.47	0.00	0.35
26–30 ns, 5	0.15	0.10	0.43	0.07	0.00

Table S14: Eigenmode cross dot products of 5 ns patches for the CL big simulation.

	1^{st} mode/2^{nd} mode				
	1	2	3	4	5
6–10 ns, 1	0.00	0.03	0.64	0.06	0.04
11–15 ns, 2	0.03	0.00	0.71	0.05	0.00
16–20 ns, 3	0.62	0.13	0.00	0.68	0.67
21–25 ns, 4	0.00	0.25	0.54	0.00	0.02
26–30 ns, 5	0.01	0.11	0.50	0.04	0.00

Table S15: Eigenmode cross dot products of 5 ns patches for the OL big simulation.

	1^{st} mode/2^{nd} mode				
	1	2	3	4	5
6–10 ns, 1	0.00	0.23	0.24	0.32	0.00
11–15 ns, 2	0.29	0.00	0.06	0.13	0.02
16–20 ns, 3	0.26	0.07	0.00	0.10	0.05
21–25 ns, 4	0.26	0.04	0.05	0.00	0.07
26–30 ns, 5	0.05	0.22	0.19	0.31	0.00

res#	CL	CL/big	CU	OU	OL	OL/big	OL/N-op[a]	OL/N-cl[b]	OL/N2-op[c]	OL/N2-cl[d]
89	99	93	96	90	0	0	27	99	67	93
181	70	57	58	42	0	0	11	96	38	59
182	99	93	96	90	0	0	27	99	67	93

[a]OL/N simulation while still open.
[b]OL/N simulation after closure.
[c]OL/N2 simulation while still open.
[d]OL/N2 simulation after closure.

Table S16: Percent probability to find anti-parallel β sheet in residues 89, 181, 182.

residue	simulation				DynDom			
	OL/N		OL/N2		OL/N		OL/N2	
	Φ	Ψ	Φ	Ψ	Φ	Ψ	Φ	Ψ
85	4	0	7	29				
86	4	1	37	2		5.8		
87	3	3	2	10	12.8	1.4		9.3
88	11	3	25	8	11.5	4.8	9.4	7.7
89	33	2	24	5	44.7	2.7	42.7	2.5
90	29	14	18	3	22.2		41.4	
180	22	2	3	2	16.2	2.9	8.7	9.8
181	15	6	1	10	9.9	4.1	0.6	10.2
182	10	8	12	5	3.4	2.9	48.7	13.9
183	15	20	5	15	14.1	19.3	11.4	10.2
184	14	1	11	4	15.8		14.5	
185	0	0	1	2				
186	1	4	8	4				

Table S17: Change in dihedrals in the hinge region from MD simulations and rigid body analysis with DynDom. MD simulation results are based on averages. DynDom results are based on the average open and the average closed structure. All values in degree.

4 Figures

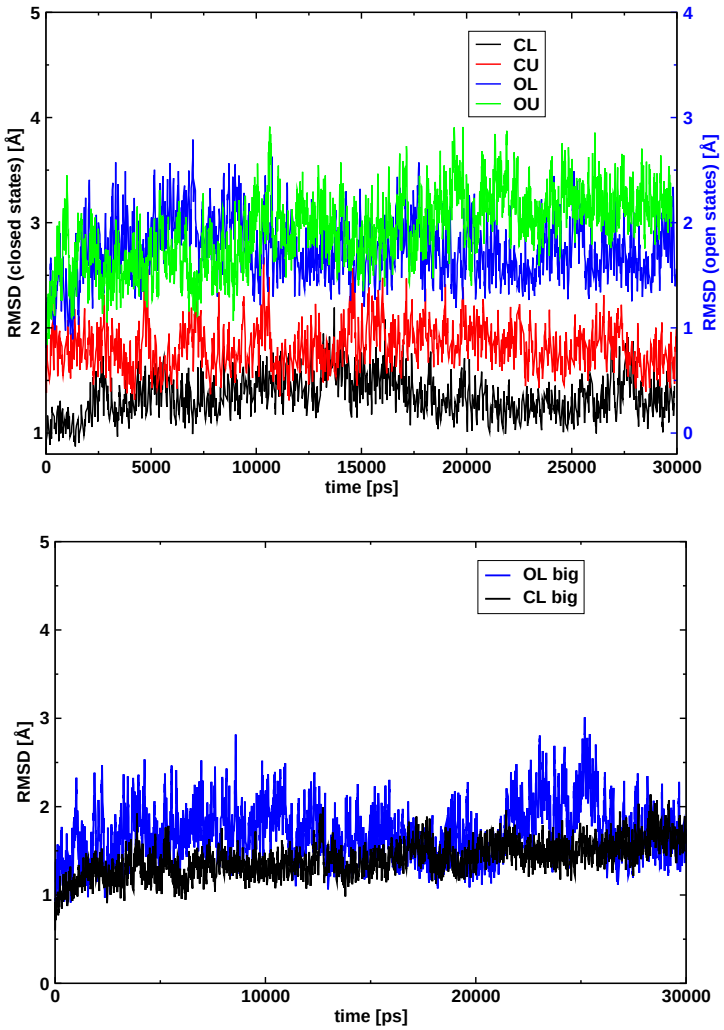


Figure S1: Root mean square deviations (RMSD) computed from all C-α atoms. Note that the open states have been shifted along the y-axis by +1 Å in the top panel as indicated by the alternative y-axis on the right of the graph. Top: smaller box sizes, bottom: larger box sizes.

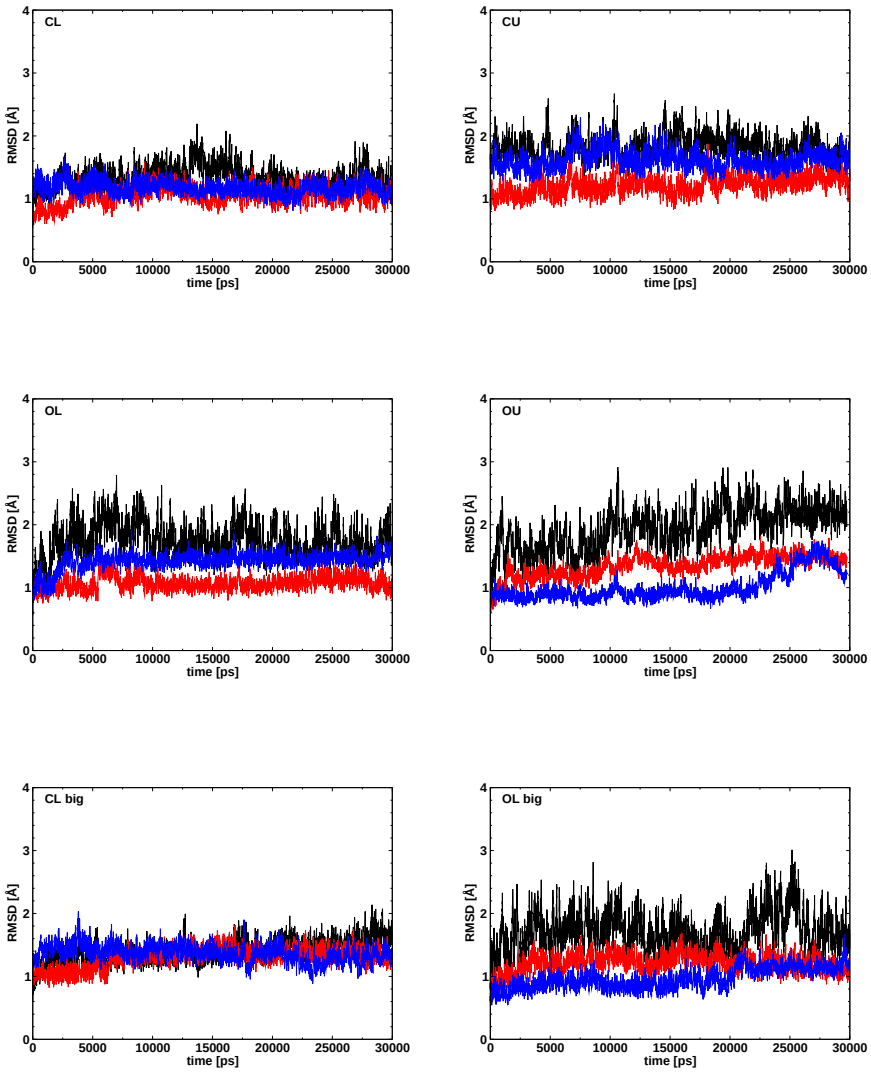


Figure S2: Comparison of RMSD fluctuations of individual domains. The reference structure is 1WDN for the closed states and 1GGG for the open states. Black: total RMS, red: large domain RMSD, blue: small domain RMS. From top left to right and down: CL closed-liganded, CU closed-unliganded, OL open-liganded, OU open-unliganded, CL big closed-liganded bigger box, OL big open-liganded bigger box.

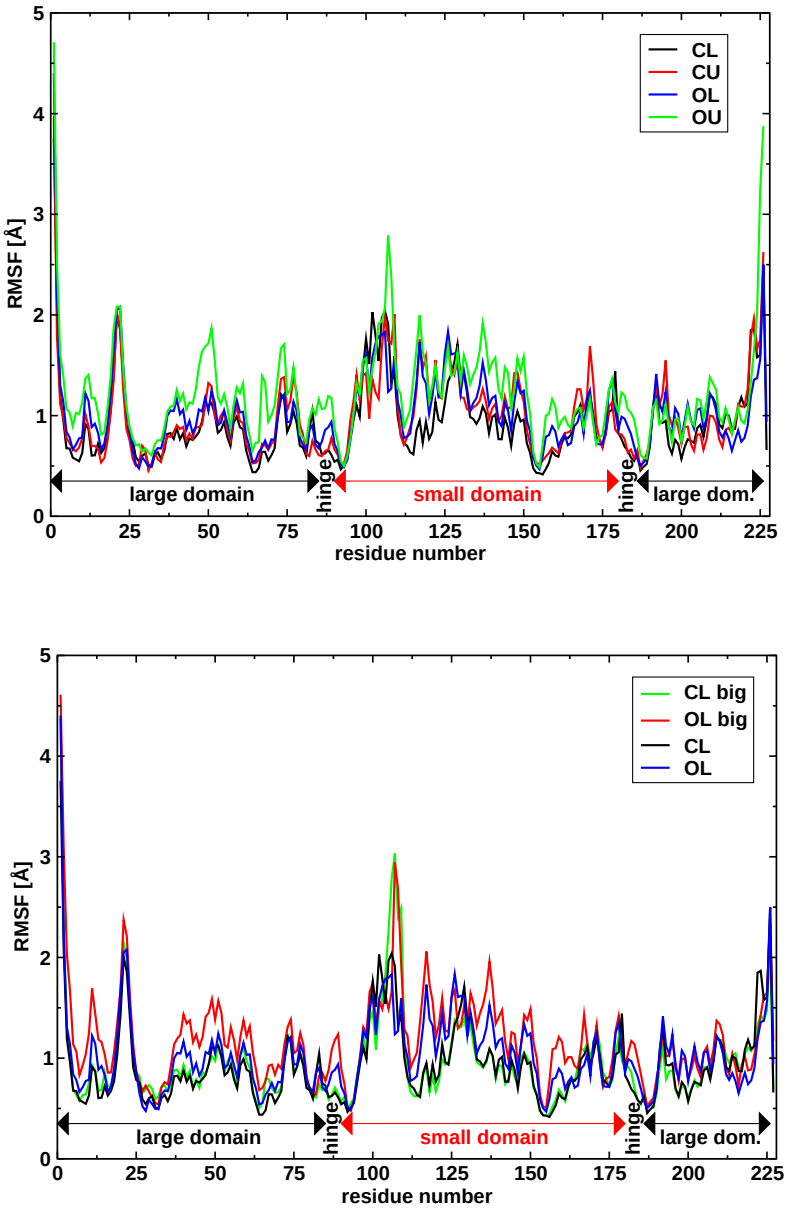
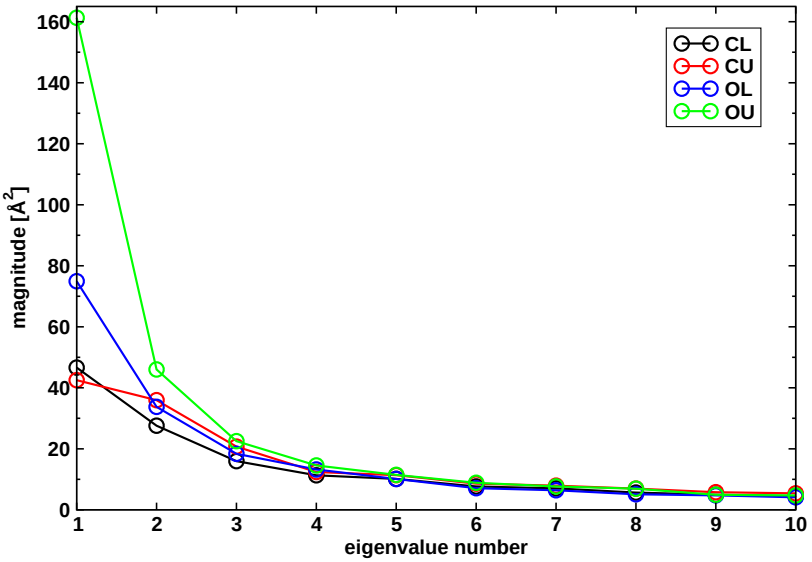


Figure S3 Per-residue root mean square fluctuations. The annotations show the residues numbers of the hinge regions and the large and small domains. The ligand is the last residue with number 227. Top: smaller box sizes, Bottom: larger box sizes in comparison with smaller box sizes.



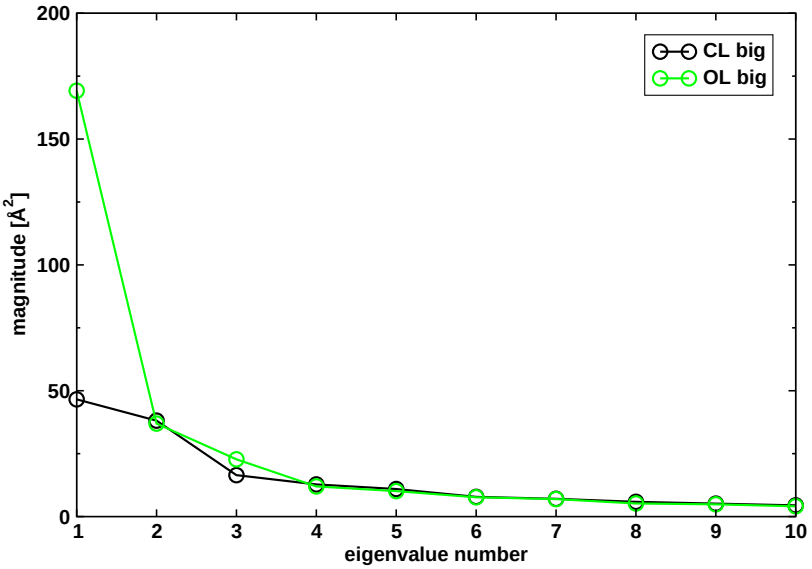


Figure S4 The magnitudes of the first ten eigenvalues calculated from the last 25 ns of each simulation. Top: smaller box sizes, Bottom: larger box sizes.

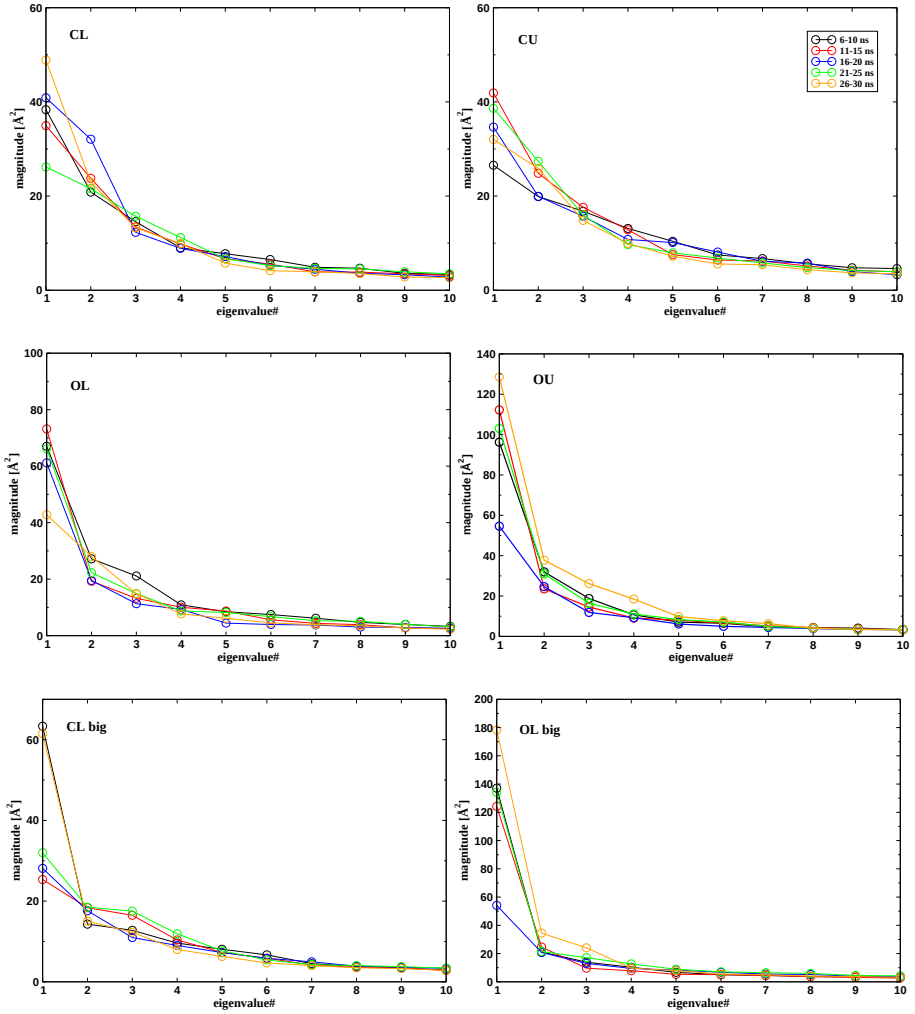


Figure S5: Magnitudes of the first ten eigenvalues calculated from five 5 ns patches. Note the different scales on the y-axes. Black: 6-10 ns, red: 11-15 ns, blue: 16-20 ns, green: 21-25 ns, orange: 26-30 ns

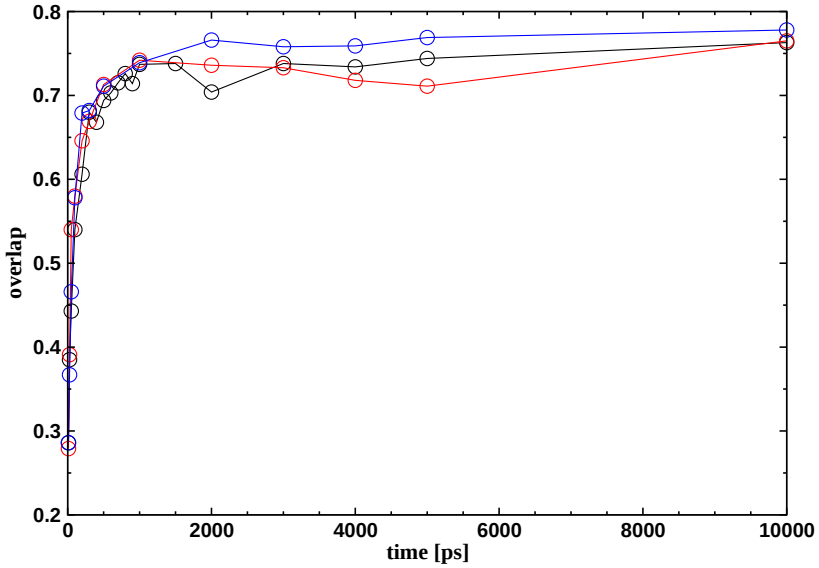


Figure S6: The overlap function estimated for selected cases calculated for the subspace spanned by the first 25 eigenvectors. The function has been calculated by comparing the eigenvectors from the initial part of the trajectory and from the parts towards the end of the trajectory. Black: CL, red: OL, blue: OL larger box

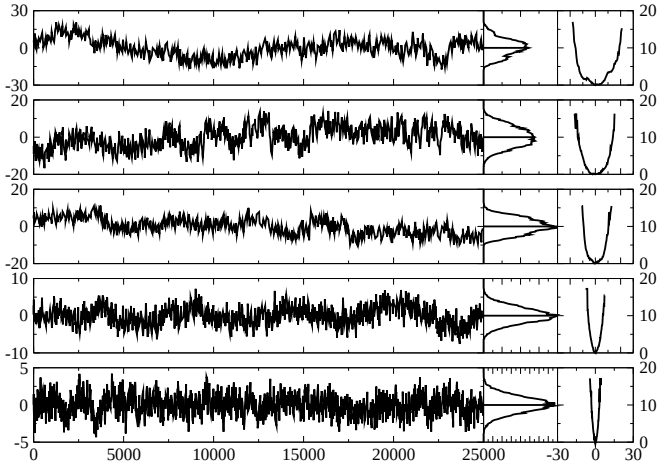


Figure S7: The projections of the first three, the 10^{th} and the 25^{th} principal component vectors onto the CL trajectory. The central panel shows the probabilities p and the right panel the free energies $G = -kT \ln p$ in kJ/mol.

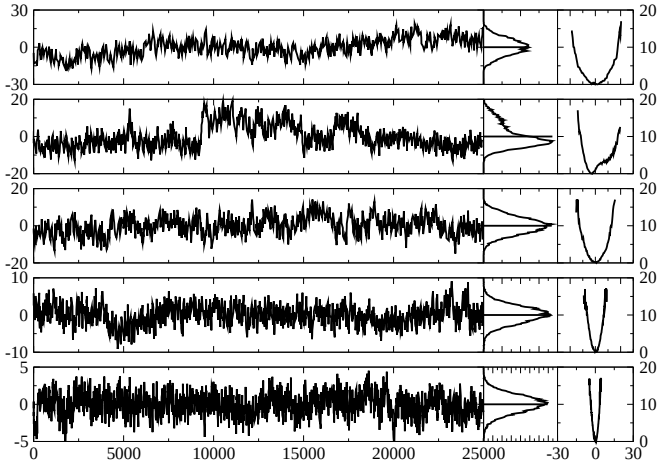


Figure S8: The projections of the first three, the 10^{th} and the 25^{th} principal component vectors onto the CU trajectory. The central panel shows the probabilities p and the right panel the free energies $G = -kT \ln p$ in kJ/mol.

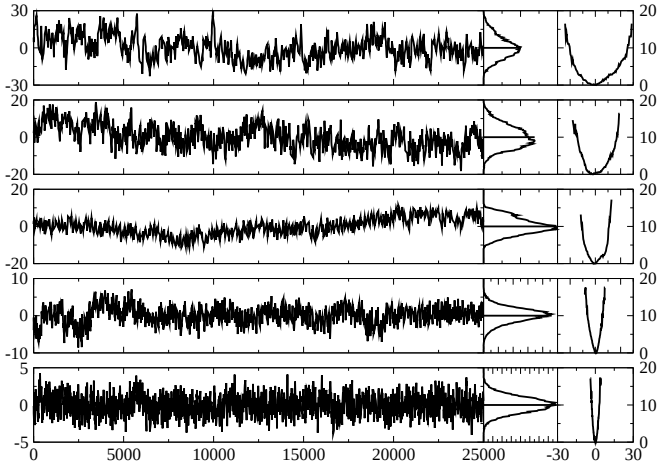


Figure S9: The projections of the first three, the 10^{th} and the 25^{th} principal component vectors onto the OL trajectory. The central panel shows the probabilities p and the right panel the free energies $G = -kT \ln p$ in kJ/mol.

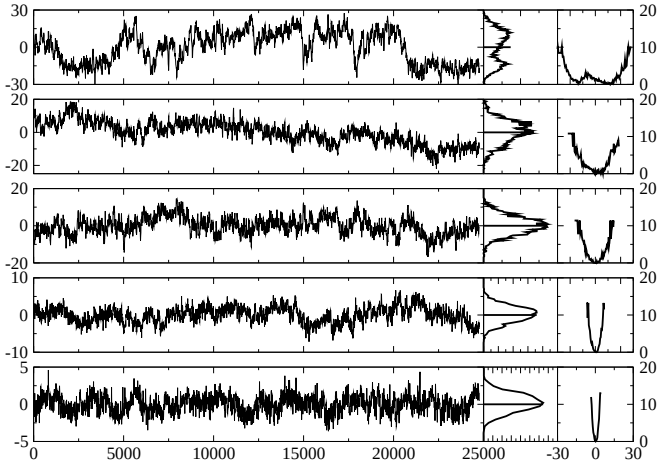


Figure S10: The projections of the first three, the 10^{th} and the 25^{th} principal component vectors onto the OU trajectory. The central panel shows the probabilities p and the right panel the free energies $G = -kT \ln p$ in kJ/mol.

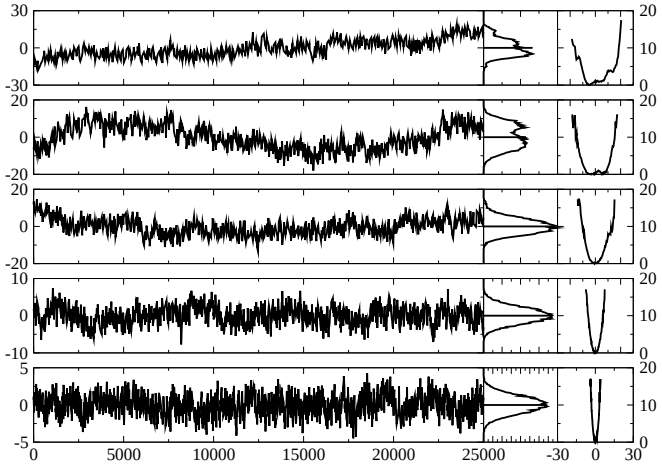


Figure S11: The projections of the first three, the 10^{th} and the 25^{th} principal component vectors onto the CL big trajectory. The central panel shows the probabilities p and the right panel the free energies $G = -kT \ln p$ in kJ/mol.

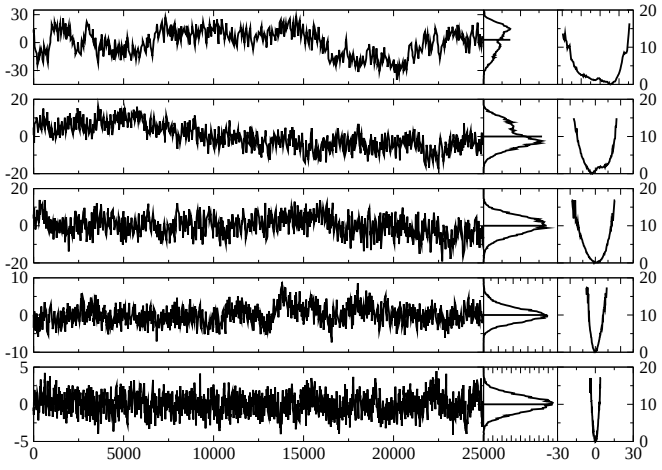


Figure S12: The projections of the first three, the 10^{th} and the 25^{th} principal component vectors onto the OL big trajectory. The central panel shows the probabilities p and the right panel the free energies $G = -kT \ln p$ in kJ/mol.

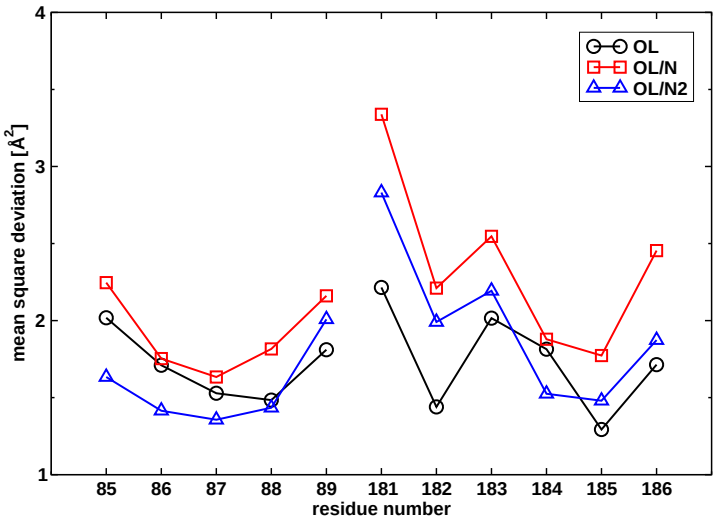


Figure S13: Mean square deviations in Ångström for the hinge region as estimated from ANM analysis. The MSDs of residues 181–186 are quite different in the OL, OL/N and OL/N2 simulations. The hinge in the OL simulations varies the least. MSDs of residues 85–89 are quite similar.

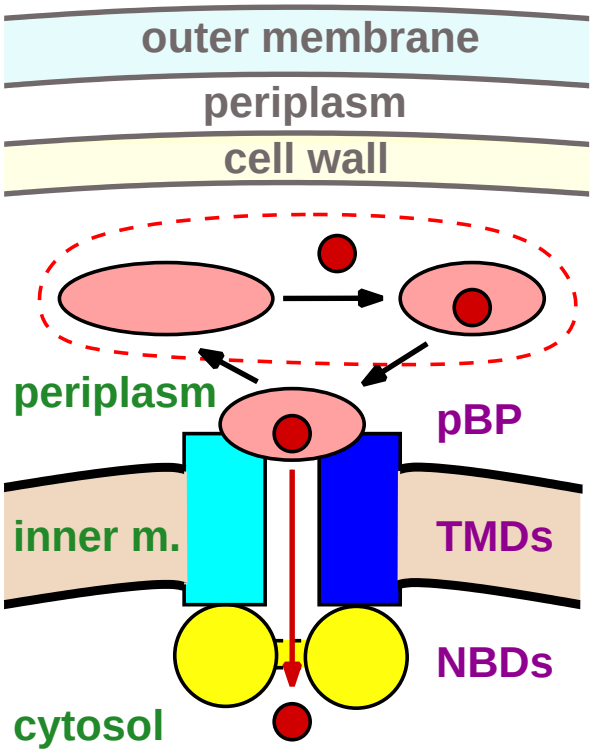


Figure S14 Schematical depiction of ABC transporter systems of Gram-negative bacteria like *E. Coli*. The broken encirclement shows the focus of this work. GlnBP (as a representative of pBPs) may undergo large conformational changes upon ligand binding. This change of shape makes it recognizable by the membrane receptor system eventually leading to the release of the ligand into the cytosol. pBP: periplasmic binding protein, TMD: trans-membrane domain, NBD: nucleotide-binding domain.

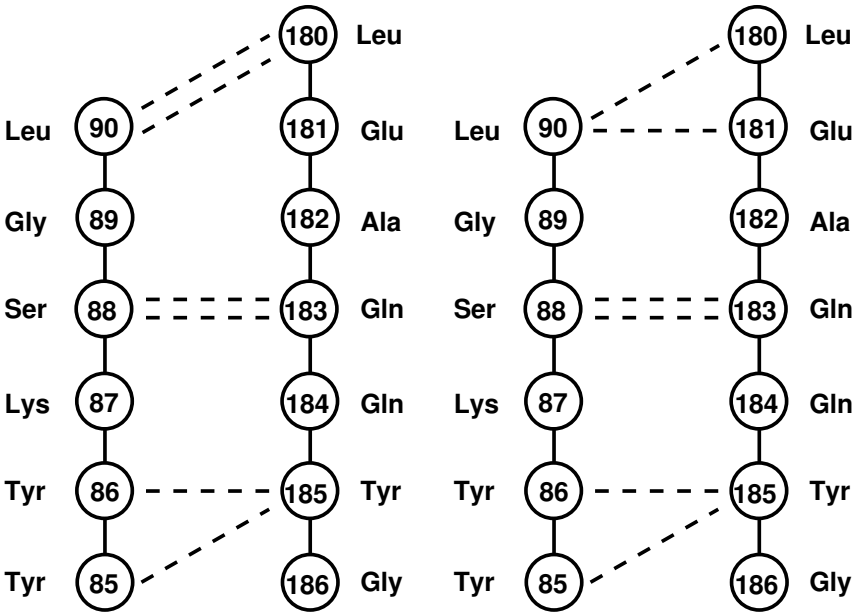


Figure S15 Hydrogen bond patterns in the backbones of the two hinge strands. The small domain is located towards the top of the graphs and the large domain towards the bottom. Left: pattern in the crystal structures 1WDN and 1GGG, and in the open state simulations, right: pattern in the closed state simulations.

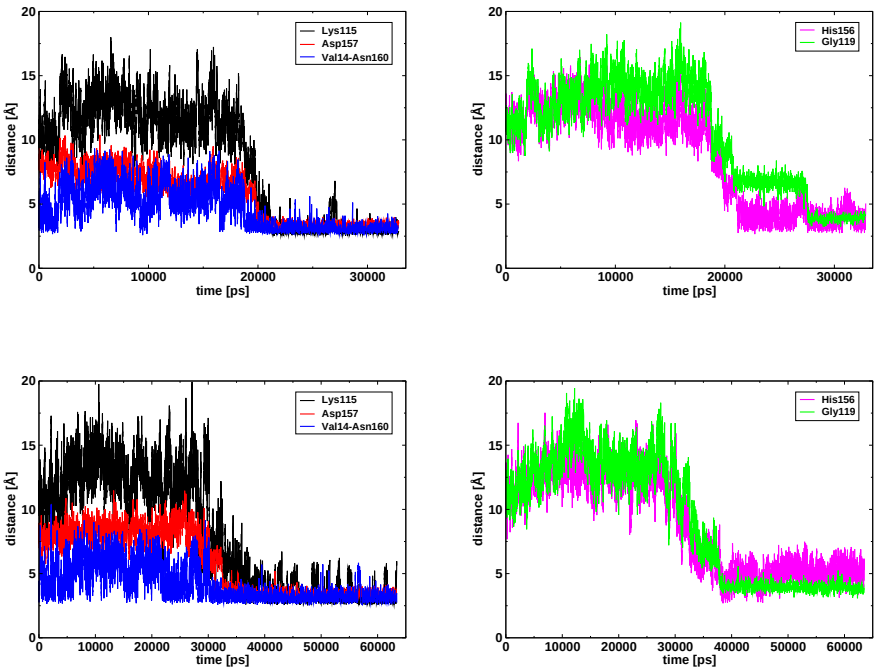


Figure S16 Formation of hydrogen bonds between Gln ligand and the residues from the small domain in the OL/N (top) and OL/N2 (bottom) simulations. The hydrogen bond between large domain residue Val 14 and small domain residue Asn 160 is also included. Shown are the distances between the heavy atoms of hydrogen bond donors and acceptors except in the case of carboxylate groups where the distance to the carboxylate carbon is plotted instead.