BIOINFORMATION
Discovery at the interface of physical and biological sciences

open access www.bioinformation.net Hypothesis

Volume 9(1)

Role of highly central residues of P-loop and it's flanking region in preserving the archetypal conformation of Walker A motif of diverse P-loop NTPases

Ekta Pathak[1], Neelam Atri[2] & Rajeev Mishra[1]*

[1]Department of Bioinformatics (MMV), Banaras Hindu University, India; [2]Department of Botany (MMV), Banaras Hindu University, Varanasi - 221005, India; Rajeev Mishra – Email: mishrarajeev@gmail.com; Phone +91 9935338891; *Corresponding author

Received December 03, 2012; Accepted December 05, 2012; Published January 09, 2013

Abstract:
P-loop NTPases represent a large and highly diverse protein family that is involved in variety of cellular functions. Walker A motif forms a typical arched conformation, necessary to accommodate the phosphate moiety of the nucleoside tri (or di-) phosphate in P-loop NTPases. The feature that maintains the ancient architecture of P-loop is unidentified and uncharacterized. Here, using a well established global network parameter, closeness centrality, we identify that Walker A and its flanking regions (N- and C-terminal) have high density of globally connected residue positions. We find that closeness centrality of these residue positions are conserved across common structural core of diverse domains of P-loop NTPase fold. Our results suggest the potential role of globally connected residues in maintaining the local conformation of P-loop.

Background:
P-loop NTPases represent a large protein family that are involved in variety of cellular functions, for example, in signal transduction, translation, protein transport and localization, signal-sequence recognition, chromosome partitioning, and membrane transport **[1-3].** Walker A also known as phosphate binding loop (P-loop) is a common feature of P-loop NTPase fold that bind nucleotide. The consensus sequence of Walker A (GXXXXGK[S/T], where X is any residue) is often used as a motif for identifying new members of this group **[4-6].** Walker A sequences are also present in many proteins that do not form P-loop, for example, peroxidases, and enzymes like α-amylase, glutamate dehydrogenase, Taq polymerase, carbonic anhydrase, binding proteins (lectin, trypsin inhibitor), proteases, and others **[7].** Here, we investigated the features that maintain the P-loop architecture by employing a well established global network parameter closeness centrality. Protein structures can be represented as a residue–residue interaction network where the residues are nodes and interactions between them constitute edges. This approach has been useful in various studies like predicting functional residues in enzyme families **[8],** protein structure flexibility **[9],** protein folding **[10],** and side-chain clusters **[11].** Closeness centrality is a global network parameter that correlates more accurately with critical residues than any other centrality measurement tested **[12].** High closeness residues interact directly or by a few intermediates with all other residues of the protein **[13].** By definition, closeness-centrality is calculated by mean distance of a node (residue) to all other nodes (residue) in the network. Amitai *et al.,* **[8]** have shown that important residue positions like those involved in substrate and co-factor binding, catalysis, and mutation intolerant residues show high closeness centrality in networks. Del sol *et al*. **[13]** have shown that centrality residues integrate and propagate the information to all other residues in protein.

Here, we show that Walker A and its flanking regions (N- and C-terminal) have high density of high closeness centrality

residue positions in P-loop NTPases. We report that closeness centrality of these residue positions are conserved across common structural core of Ras superfamily and diverse domains of P-loop NTPase fold. No such high densities of high centrality residue positions are observed in the proteins containing Walker A sequence that do not form P-loop. The presented data clearly indicate the role of globally connected residues in conservation of the local conformation of an ancient motif such as Walker A.

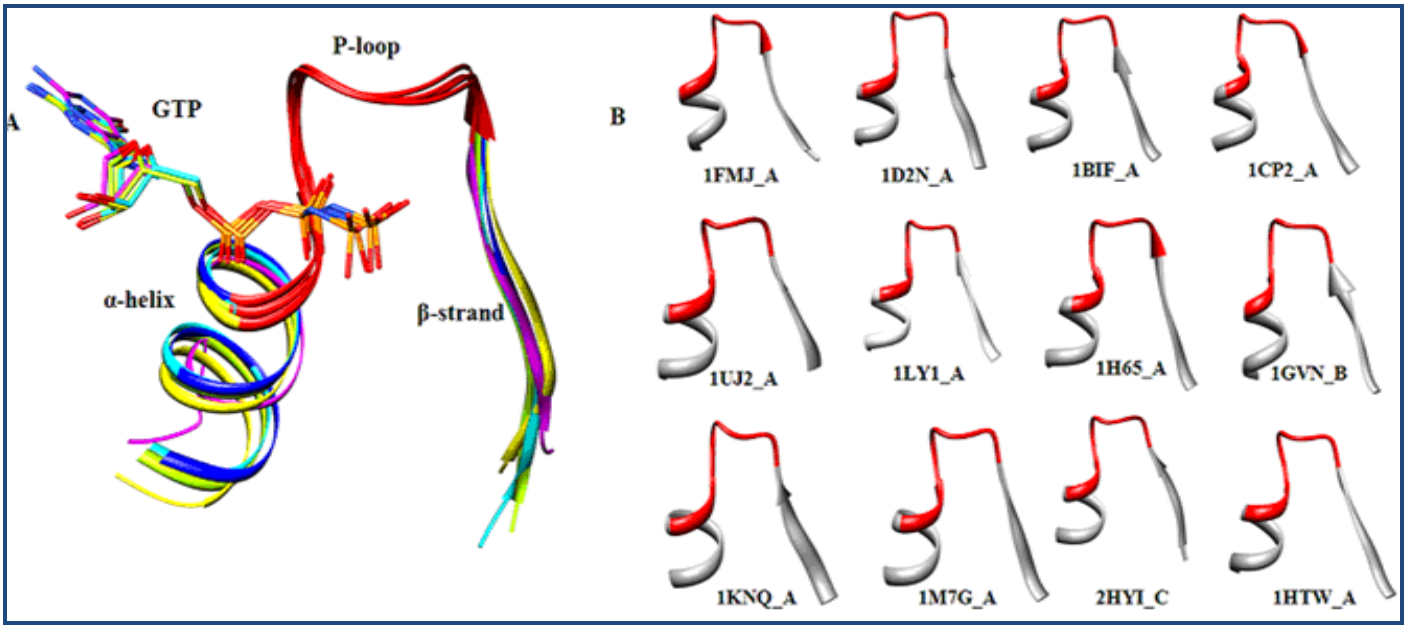

Figure 1: A) Ribbon diagram of typical architecture of P-loop (Red) with bound nucleotide molecule (stick) of Ras superfamily proteins [Ras (green), Rab (cyan), Rho (blue), Ran (yellow), and Arf (magenta)]; **B)** Ribbon diagrams of typical architecture of P-loop (red) in representatives of diverse P-loop containing NTPases. 4 letter words are the PDBID**.**

Methodology:

Selection of structures of P-loop containing NTPases

High resolution X-ray crystallographic structures of diverse domain of P-loop containing NTPases were used in the study. Initially, ScopTree search of protein databank (http://www.rcsb.org/pdb) was used to retrieve a set of 1203 structures of P-loop containing nucleoside triphosphate hydrolase. The search was then refined to 227 distantly related protein structures by using ScopTree homologue removal tool at 30% sequence identity cutoff. This was primarily done to avoid redundancy and utilize the diversity present in the P loop NTPases. Complete structures (i.e., without chain breaks or missing residues) with resolution ≤ 2.4 were chosen. Finally, we selected 23 structures of P-loop NTPases **Table 1 (see supplementary material).** We retrieved 22 PDB files for protein structures containing Walker A sequence (GXXXXGKS/T) that do not form the P-loop **Table2 (see supplementary material) [7].**

Computation of closeness centrality

Protein structures can be represented as a residue–residue interaction graphs in which amino acid residues serve as the nodes and their interatomic contacts are the edges. Closeness centrality correlates more accurately with critical residues than any other centrality measurement tested **[12]**. Therefore, we used SARIG server which efficiently calculates the closeness centrality **(please see supplementary material for calculation and explanation).**

Beginning with the atomic coordinates of a protein structure, server calculates the interaction between each pair of atoms by using the CSU program **[14].** Closeness values were calculated for each residue and standardized by calculating their standard deviation from the mean value. The z-score of the closeness centrality was calculated by z-score = $(C(x) - \mu) / \sigma$, where μ is the mean value of closeness and σ is the standard deviation. The residues with z-score ≥ 1.0 were considered significant (for detailed descriptions, please refer to Amitai *et al* **[8]).** Protein structure analysis was performed using Chimera (http://plato.cgl.ucsf.edu/chimera).

Results and Discussion:

Walker A motif forms a typical architecture in P-loop fold NTPase **(Figure1A & 1B).** A distortion in the P-loop conformation makes it incompatible with the binding of nucleotides **[15].** The features that contribute in preserving the architecture of this ancient motif remain unidentified and uncharacterized. Therefore, an important and open question is how P-loop forms a typical architecture in structurally and functionally diverse P-loop NTPases. Here, we used a well established closeness centrality network parameter to study the global impact of residues on the typical local conformation of P-loop. Residues with high closeness value are central in network and interact with other residues directly or by a few intermediates **[8].**

High closeness residue positions around P-loop and its flanking regions in Ras Super family members

In order to understand the P-loop architecture, we first analyzed the residue-residue interaction network of Ras superfamily (Ras: 5P21; Rab: 3RAB; Ran: 1IBR, Rho: 1M7B and Arf: 1R4A) experimental structures in GTP bound form. Interestingly, Walker A and its flanking regions showed high density of high closeness residue positions **(Table 1).** Here, the

high closeness centrality positions are defined as those positions with statistically significant closeness values (z-score ≥1.0). Five Walker A residue positions (W1, W2, W5, W6, W7), four contiguous N-terminal residue positions (N2-N5) and two C-terminal residues (C2 and C3), flanking the Walker A, showed high closeness centrality in Ras superfamily members **(Table 1)**.

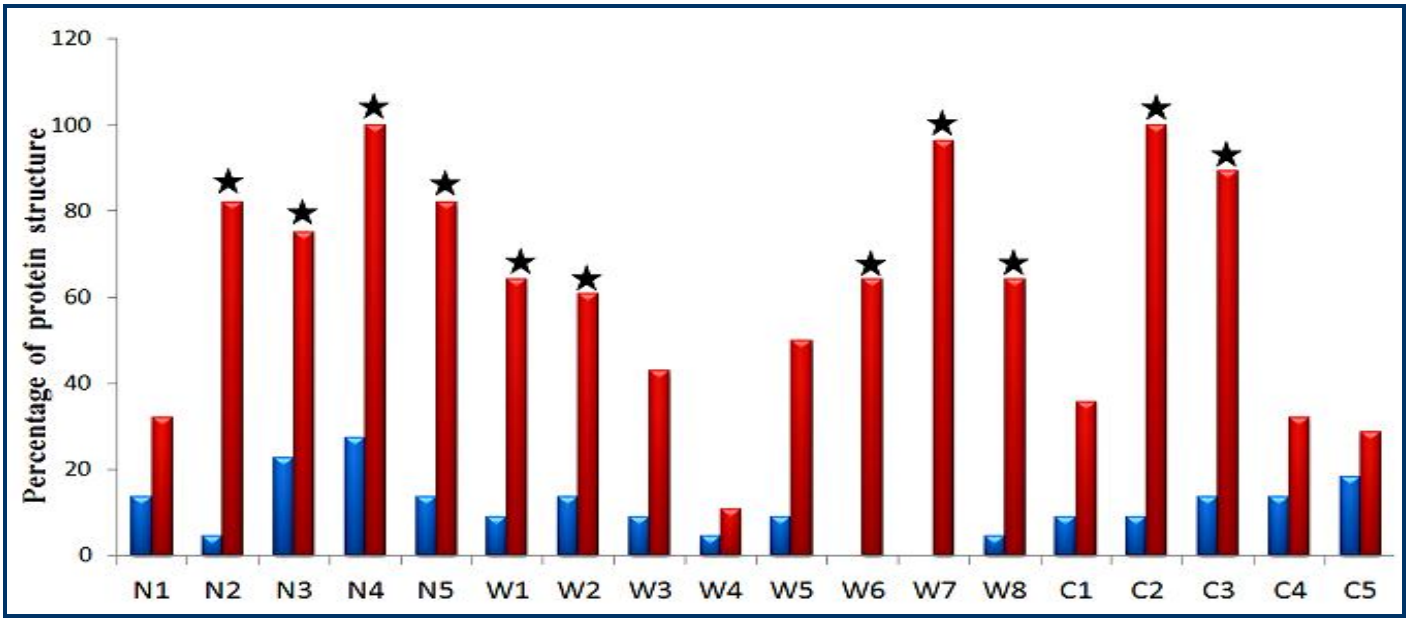

Figure 2: Bar graph showing distribution of high closeness residue positions in diverse set of P-loop NTPases (red bar). Walker A containing proteins that do not form P-loop are depicted in blue bar. High density of high closeness residue positions (star marked) shown around Walker A (W1-W8) and its flanking regions N terminal (N1-N5) and C terminal (C1-C5).

High density of high closeness residue positions in P-loop and its flanking regions in diverse set of P-loop NTPases

Since the Ras superfamily belongs to P-loop NTPase fold, we then extended the centrality analysis on high resolution X-ray crystallographic structures of P-loop NTPases **(Table 1).** The structural overlay of highly diverse P-loop NTPases fold showed that the typical P-loop architecture is maintained **(Figure 1B).** In order to avoid redundancy and utilize the diversity present in the P loop NTPases, we selected a set of 23 NTPase structures at 30% sequence identity cutoff **(see methodology)**. We wanted to look at the impact of sequence diversity on the closeness value of the residues of P-loop and its flanking region. Intriguingly, the highly diverse P-loop NTPases exhibited a similar pattern of high density of conserved high closeness centrality residue positions around Walker A motif, as seen in Ras Super family. Here, the conserved high closeness centrality positions are defined as those positions with statistically significant closeness values (z-score ≥1.0) in at least 60% of the structures of P-loop NTPase fold **(Figure 2 & Table 1).** 11 such residue positions around Walker A and its flanking regions showed high closeness value. Four contiguous residue positions (N2-N5) of the N-terminal, two residue positions of C-terminal (C2-C3) and five residue positions of Walker A (W1, W2, W6, W7, W8) were showing high closeness centrality. The residue positions N4 (100%), W7 (96%) and C2 (100 %) were highly conserved in their centrality across the diverse structures. The invariant residue positions (G, K, S/T) and variant residue positions (W2) of Walker A showed high closeness centrality **(Table 1).** Walker A sequence has wider distribution and observed in many proteins that do not bind nucleotides **[7].** The structural analysis revealed that these proteins do not form the conspicuous P-loop architecture **[7].** To test our prediction, we calculated the closeness value in Walker A sequences that do not form P-loop **(Table 2)**. We did not observe high density of high closeness centrality pattern.

Our results indicate the high density of conserved high closeness residue positions in P-loop and its flanking regions in P-loop fold NTPase and underscore its role in supporting the architecture of P-loop. The study presented is in concord with the observation that highly central residue positions correlate well with active site residues or their neighbors that provide supportive scaffold **[13].** However, high closeness value of invariant (G, K, and S/T) residues of Walker A indicates its role in catalysis. P-loop lysine interacts and forms hydrogen-bond with oxygen of γ -phosphate of bound nucleotide and serine/threonine binds with Mg^{2+} **[16, 17].** Recently Grüber *et al*. **[15]** demonstrated the role of conserved glycine residues of Walker A motif in guarding the active-site region for nucleotide entrance in archaea-type ATP synthases. The altered conformation of the P- loop resulted in the active-site region being closed to nucleotide entry **[15].**

Conclusion:

In the context of network, protein structural scaffold and sequence diversity can be visualized as a dramatic change in the type of node, and also the connections between the nodes. Regardless of such diversity, depicted in Ras superfamily and diverse domains of P-loop fold NTPase, the closeness centrality of residue positions in P-loop and its flanking regions are remarkably maintained to be high. Thus, our finding supports the observation that centrality of a residue is maintained evolutionarily to assure the proper functioning of protein **[8, 13].** We did not find such high centrality residue positions in proteins containing Walker A motif that do not form P-loop. This strengthens the evidence that required geometry of

archetypal P-loop is achieved by high density of residue positions which are globally connected in short steps.

Acknowledgment:
EP acknowledges CSIR, India for the JRF and SRF. We thank Department of Bioinformatics, MMV, Banaras Hindu University, for providing computational facility.

References:

[1] Koonin EV *et al. Trends Biochem Sci*. 2000 **25**: 223 [PMID: 10782090]
[2] Saraste M *et al. Trends Biochem Sci.* 1990 **15**: 430 [PMID: 2126155]
[3] Vetter IR *et al. Q Rev Biophys*. 1999 **32**: 1 [PMID: 10800520]
[4] Walker JE *et al. EMBO* J. 1982 **1:** 945 [PMID: 6329717]
[5] Wolf YI *et al. Genome Res*. 1999 **9**: 17 [PMID: 9927481]
[6] Ma BG *et al. Biochem Biophys Res Commun*. 2008 **366**: 607 [PMID: 18073136]
[7] Ramakrishnan C *et al. Protein Eng*. 2002 **15**: 783 [PMID: 12468712]
[8] Amitai G *et al. J Mol Biol.* 2004 **344**: 1135 [PMID: 15544817]
[9] Jacobs *DJ et al. Proteins*. 2001 **44:** 150 PMID: 11391777]
[10] Vendruscolo M *et al. Phys Rev E Stat Nonlin Soft Matter Phys*. 2002 **65**: 0619101 [PMID: 12188762]
[11] Kannan N & Vishveshwara S, *J Mol Biol*. 1999 **292:** 441 [PMID: 10493887]
[12] Thibert B *et al. BMC Bioinformatics*. 2005 **6:** 213 [PMID: 16124876]
[13] del Sol A *et al. Protein Sci*. 2006 **15**: 2120 [PMID: 16882992]
[14] Sobolev V *et al. Bioinformatics*. 1999 **15**: 327 [PMID: 10320401]
[15] Kumar A *et al. J Mol Biol*. 2010 **396**: 301 [PMID: 19944110]
[16] Crampton DJ *et al. Biochemistry*. 2001 **40**: 3710 [PMID: 11297439]
[17] Runquist JA *et al. Biochemistry*. 2001 **40**: 14530 [PMID: 11724566]

Edited by P Kangueane
Citation: Pathak *et al*. Bioinformation 9(1): 023-028 (2013)

BIOINFORMATION

open access

Supplementary material:

Methodology:

Computation of closeness centrality

Protein structures can be represented as a residue–residue interaction graphs in which amino acid residues serve as the nodes and their interatomic contacts are the edges. Closeness centrality correlates more accurately with critical residues than any other centrality measurement tested **[12].** Therefore, we used SARIG server which efficiently calculates the closeness centrality (http://bioinfo2.weizmann.ac.il/~pietro/SARIG/V3/index.html). Closeness centrality of node x(C (x)) is calculated as follows:

$$C(x) = (N-1)/\sum_{y \epsilon U,\, y \neq x} d(x, y)$$

Where d(x, y) is the shortest-path between node x and any node y. U is the set of all nodes and N is the number of nodes in the network.

Table 1: Closeness centrality (z-scores) of residues of Walker A and flanking region sequence in P-loop NTPases H (%) is the percentage of structures with z- score ≥ 1.0. HCR is the number of High closeness residues with z –score ≥ 1.0.

PDB ID	N1	N2	N3	N4	N5	W1	W2	W3	W4	W5	W6	W7	W8	C1	C2	C3	C4	C5	HCR
1FMJ_A	ASP	VAL	PHE	VAL	ALA	SER	TYR	GLN	ARG	SER	GLY	THR	THR	MET	THR	GLN	GLU	LEU	12
	-0.147	0.819	1.025	1.712	1.696	1.688	2.342	2.36	2.287	1.919	1.467	1.736	1.413	0.928	1.132	0.935	0.409	0.433	
1D2N_A	SER	VAL	LEU	LEU	GLU	GLY	PRO	PRO	HIS	SER	GLY	LYS	THR	ALA	LEU	ALA	ALA	LYS	15
	0.582	1.533	1.423	1.852	1.666	1.245	1.685	1.151	0.895	1.637	1.618	1.954	1.704	0.968	1.637	1.647	1.378	1.387	
1NJF_A	ALA	TYR	LEU	PHE	SER	GLY	THR	ARG	GLY	VAL	GLY	LYS	THR	SER	ILE	ALA	ARG	LEU	13
	0.224	1.471	0.911	1.782	0.774	0.665	1.108	1.191	0.799	1.154	1.248	1.902	2.024	1.532	2.058	1.771	2.426	1.708	
1SVM_A	TYR	TRP	LEU	PHE	LYS	GLY	PRO	ILE	ASP	SER	GLY	LYS	THR	THR	LEU	ALA	ALA	ALA	9
	1.149	1.344	0.766	1.349	0.388	-0.155	-0.248	-0.793	-0.63	-0.169	0.548	0.938	1.078	1.243	1.733	1.442	1.526	2.014	
1BIF_A	LEU	ILE	VAL	MET	VAL	GLY	LEU	PRO	ALA	ARG	GLY	LYS	THR	TYR	ILE	SER	LYS	LYS	10
	0.412	1.164	0.368	1.277	0.436	0.051	0.304	-0.428	0.048	0.334	1.007	1.107	1.414	1.869	1.917	1.193	1.976	2.772	
1CP2_A	GLN	VAL	ALA	ILE	TYR	GLY	LYS	GLY	GLY	ILE	GLY	LYS	SER	THR	THR	THR	GLN	ASN	15
	0.695	1.329	1.384	2.396	2.114	1.644	0.882	1.086	0.18	1.157	1.06	2.466	1.615	1.366	2.202	1.735	1.46	1.183	
1G3Q_A	ILE	SER	ILE	VAL	SER	GLY	LYS	GLY	GLY	THR	GLY	LYS	THR	THR	VAL	THR	ALA	ASN	14
	1.899	1.548	2.349	1.846	1.316	-0.439	0.773	0.317	-0.317	1.186	1.082	2.468	1.859	1.461	2.233	2.077	1.244	1.232	
1NP6_B	LEU	LEU	ALA	PHE	ALA	ALA	TRP	SER	GLY	THR	GLY	LYS	THR	THR	LEU	LEU	LYS	LYS	9
	1.192	1.816	1.08	2.075	1.08	0.318	0.478	0.212	-0.42	0.55	0.8	1.658	1.192	0.571	1.154	1.658	0.582	0.603	
1YRB_A	ILE	VAL	VAL	PHE	VAL	GLY	THR	ALA	GLY	SER	GLY	LYS	THR	THR	LEU	THR	GLY	GLU	12
	1.199	1.706	1.863	2.057	2.12	1.241	1.276	0.686	0.302	0.894	1.089	1.863	1.039	0.583	1.726	1.454	0.322	0.656	
2HYI_C	ASP	VAL	ILE	ALA	GLN	SER	GLN	SER	GLY	THR	GLY	LYS	THR	ALA	THR	PHE	SER	ILE	12
	-0.194	0.593	0.331	1.475	1.418	1.412	1.469	2.001	0.847	1.429	1.257	2.333	1.97	1.139	1.192	1.807	0.902	0.393	
2J0S_A	ASP	VAL	ILE	ALA	GLN	SER	GLN	SER	GLY	THR	GLY	LYS	THR	ALA	THR	PHE	SER	ILE	12
	-0.218	0.567	0.301	1.468	1.317	1.22	1.38	2.052	0.912	1.409	1.363	2.558	2.052	0.876	1.486	2.072	1.007	0.669	
1LY1_A	ILE	ILE	LEU	THR	ILE	GLY	CYS	PRO	GLY	SER	GLY	LYS	SER	THR	TRP	ALA	ARG	GLU	11
	0.993	1.664	1.416	2.117	1.567	0.702	1.01	1.027	-0.15	0.64	1.01	2.16	1.253	0.733	1.491	1.235	-0.15	-0.2	
1VHT_A	ILE	VAL	ALA	LEU	THR	GLY	GLY	ILE	GLY	SER	GLY	LYS	SER	THR	VAL	ALA	ASN	ALA	10
	1.035	1.425	1.396	1.85	1.721	1.253	0.589	0.566	-0.072	0.628	0.788	1.689	1.115	0.574	1.3	1.377	0.74	0.387	
1H65_A	LEU	THR	ILE	LEU	VAL	GLY	LYS	GLY	GLY	VAL	GLY	LYS	SER	SER	THR	VAL	ASN	SER	9
	0.975	1.059	1.997	1.985	1.815	1.202	1.577	0.035	0.259	0.893	0.985	1.738	0.409	0.717	2.197	1.37	0.657	0.994	
1SVS_A	LYS	LEU	LEU	LEU	LEU	GLY	ALA	GLY	GLU	SER	GLY	LYS	SER	THR	ILE	VAL	LYS	GLN	14
	-0.603	0.127	0.397	1.083	0.704	1.056	2.117	1.344	1.916	1.667	1.552	1.613	1.042	1.499	1.691	1.315	1.159	1.002	
2AKA_A	SER	LEU	LEU	ILE	THR	GLY	GLU	SER	GLY	ALA	GLY	LYS	THR	GLU	ASN	THR	LYS	LYS	15
	1.363	1.46	1.735	1.808	1.96	1.255	1.735	1.393	0.903	1.649	1.516	1.921	1.31	1.187	1.718	1.436	0.717	0.855	
2AKA_B	GLN	ILE	ALA	VAL	VAL	GLY	GLY	GLN	SER	ALA	GLY	LYS	SER	SER	VAL	LEU	GLU	ASN	15
	1.38	1.901	1.638	2.2	1.79	1.372	0.795	0.729	0.154	1.074	1.248	2.053	1.568	1.121	2.28	2.311	1.414	1.524	
2CXX_A	THR	ILE	ILE	PHE	ALA	GLY	ARG	SER	ASN	VAL	GLY	LYS	SER	THR	LEU	ILE	TYR	ARG	8
	0.079	0.978	1.739	2.08	1.888	1.576	1.72	0.34	0.029	1.436	0.824	2.022	0.104	0.505	2.179	0.9	0.104	0.233	
1GVN_B	ALA	PHE	LEU	LEU	GLY	GLY	GLN	PRO	GLY	SER	GLY	LYS	THR	SER	LEU	ARG	SER	ALA	14
	0.569	1.189	1.222	2.165	1.495	1.495	1.785	1.957	1.469	1.539	1.099	2.277	0.964	1.408	1.711	1.909	0.894	0.848	
1HTW_A	MET	VAL	TYR	LEU	ASN	GLY	ASP	LEU	GLY	ALA	GLY	LYS	THR	THR	LEU	THR	ARG	GLY	8
	1.342	2.316	1.837	2.462	1.264	-0.295	-0.424	-0.347	-0.866	-0.106	0.459	1.264	0.722	0.333	1.751	1.687	0.621	0.364	
1KNQ_A	ILE	TYR	VAL	LEU	MET	GLY	VAL	SER	GLY	SER	GLY	LYS	SER	ALA	VAL	ALA	SER	GLU	10
	0.501	1.305	1.58	2.477	2.146	1.492	1.545	1.206	0.286	0.846	0.772	2.288	0.891	0.713	1.423	1.271	-0.11	-0.248	
1M7G_A	THR	ILE	TRP	LEU	THR	GLY	LEU	SER	ALA	SER	GLY	LYS	SER	THR	LEU	ALA	VAL	GLU	8
	1.498	2.106	2.474	2.294	1.898	0.893	0.935	0.425	-0.477	0.388	0.491	1.7	0.472	0.388	1.365	1.156	-0.145	0.189	
1UJ2_A	LEU	ILE	GLY	VAL	SER	GLY	GLY	THR	ALA	SER	GLY	LYS	SER	SER	VAL	CYS	ALA	LYS	9
	0.56	1.402	0.985	2.13	1.661	0.997	1.35	1.55	0.337	0.739	0.75	2.146	1.058	-0.001	1.337	1.591	0.164	-0.048	
5P21_A	LYS	LEU	VAL	VAL	VAL	GLY	ALA	GLY	GLY	VAL	GLY	LYS	SER	ALA	LEU	THR	ILE	GLN	9
	0.271	1.301	1.376	2.081	1.827	1.21	0.964	-0.37	0.032	0.846	1.068	1.953	0.683	0.242	2.235	1.192	-0.116	0.51	
3RAB_A	LYS	ILE	LEU	ILE	ILE	GLY	ASN	SER	SER	VAL	GLY	LYS	THR	SER	PHE	LEU	PHE	ARG	12
	0.191	1.363	1.326	2.011	1.805	1.29	1.095	-0.248	-0.015	1.078	1.13	1.866	1.112	0.362	2.496	1.326	-0.133	0.666	
1M7B_A	LYS	ILE	VAL	VAL	VAL	GLY	ASP	SER	GLN	CYS	GLY	LYS	THR	ALA	LEU	LEU	HIS	VAL	11
	0.258	1.175	1.054	1.978	1.764	1.099	1.411	0.345	0.616	1.427	1.039	2.033	0.603	0.511	2.28	1.331	-0.169	0.577	
1IBR_A	LYS	LEU	VAL	LEU	VAL	GLY	ASP	GLY	GLY	THR	GLY	LYS	THR	THR	PHE	VAL	LYS	ARG	9
	0.642	1.45	1.34	1.976	1.854	1.129	0.977	-0.251	0.246	0.977	1.216	2.059	0.782	0.782	2.429	1.487	0.029	0.551	
1R4A_A	ARG	ILE	LEU	ILE	LEU	GLY	LEU	ASP	GLY	ALA	GLY	LYS	THR	THR	ILE	LEU	TYR	ARG	8
	0.124	1.192	1.23	1.802	1.716	0.968	1.467	-0.138	-0.475	1.078	0.485	1.653	0.035	-0.009	1.868	0.95	0.468	0.246	
H (%)	32.14	82.14	75.00	100	82.14	64.29	60.71	42.86	10.71	50.00	64.29	96.42	64.29	35.71	100	89.29	32.14	28.57	

Table 2: Closeness centrality (z-score) of residues of proteins containing Walker A sequence, which do not form P-loop. H (%) is the percentage of structures with z- score ≥ 1.0

PDBID	N1	N2	N3	N4	N5	W1	W2	W3	W4	W5	W6	W7	W8	C1	C2	C3	C4	C5
6MHT	PHE	ALA	LYS	THR	GLY	GLY	TYR	LEU	VAL	ASN	GLY	LYS	THR	ARG	LYS	LEU	HIS	PRO
_A	-	-	1.051	0.834	0.091	-	0.759	-	-	-	-1.55	-	-	-	0.07	0.79	0.201	0.33
	0.036	0.011				0.387		0.564	0.494	1.454		0.619	0.342	0.036				
1QPA_A	GLU	THR	GLN	LEU	ARG	GLY	THR	ALA	PHE	PRO	GLY	LYS	THR	GLY	ILE	GLN	GLY	THR
	0.349	0.462	0.037	-	-0.83	-	-1.43	-1.1	-	0.027	-	-	-	-	-	-0.14	0.169	0.271
				0.155		0.789			0.908		0.696	0.675	1.518	0.975	0.314			
2CYP_A	VAL	VAL	ALA	LEU	MET	GLY	ALA	HIS	ALA	LEU	GLY	LYS	THR	HIS	LEU	LYS	ASN	SER
	0.889	0.293	1.175	1.736	1.388	1.254	1.874	1.072	0.776	1.834	0.642	-	-	0.01	-	-	0.596	0.72
												0.063	0.249		0.916	0.384		
1DS4_A	VAL	VAL	ALA	LEU	MET	GLY	ALA	GLY	ALA	LEU	GLY	LYS	THR	HIS	LEU	LYS	ASN	SER
	0.799	0.185	1.136	1.779	1.368	0.972	1.91	0.402	0.694	1.951	0.782	0.585	-	0.263	-	-0.37	0.616	0.616
													0.289		0.791			
1MTY_D	ILE	ALA	GLU	LEU	HIS	GLY	LEU	ARG	SER	ASP	GLY	LYS	THR	LEU	ILE	ALA	GLN	PRO
	-	-	-	0.079	0.012	-	-0.04	0.186	-	-	-	-	-	0.001	0.739	0.431	0.591	0.014
	0.228	0.654	0.639			0.115			0.959	0.901	0.806	0.824	0.292					
1MMO_	ILE	ALA	GLU	LEU	HIS	GLY	LEU	ARG	SER	ASP	GLY	LYS	THR	LEU	ILE	ALA	GLN	PRO
D	-	-	-	0.066	-	-	-	0.19	-0.97	-0.96	-	-	-	-	0.711	0.393	0.614	-
	0.244	0.697	0.667		0.023	0.135	0.059				0.805	0.871	0.384	0.044				0.031
1STE_A	SER	LYS	ASP	ASN	VAL	GLY	LYS	VAL	THR	GLY	GLY	LYS	THR	CYS	MET	TYR	GLY	GLY
	-	-	-	-	-	-	-	-	-	-	-	-	0.457	0.431	1.583	1.629	1.606	1.875
	1.242	1.294	1.386	1.361	2.039	2.054	0.989	1.998	1.533	0.926	1.102	0.155						
1QSS_A	LEU	PHE	ASP	GLU	LEU	GLY	LEU	PRO	ALA	ILE	GLY	LYS	THR	GLU	LYS	THR	GLY	LYS
	-	-	-	-	-1.82	-	-	-1.62	-	-	-	-	-	-2.17	-	-	-	-1.53
	1.454	1.224	1.879	2.083		1.898	1.734		1.611	1.113	1.644	1.934	1.544		2.658	2.172	2.143	
1WGT_	THR	SER	LYS	ARG	CYS	GLY	SER	GLN	ALA	GLY	GLY	LYS	THR	CYS	PRO	ASN	ASN	HIS
A	-	0.132	0.721	0.934	0.531	0.449	-	0.439	-	-	0.225	0.9	1.925	1.751	1.477	1.184	2.247	1.626
	0.508						0.038		0.409	0.485								
1CWV_	ALA	VAL	ILE	GLY	ASP	GLY	ALA	PRO	ALA	ASN	GLY	LYS	THR	ALA	ILE	THR	VAL	GLU
A	-	-	-	-	-1.06	-	-	-	-0.47	-0.5	-	-	-	-	-	-	-	-
	1.315	1.311	1.045	1.065		0.771	0.763	0.461			0.803	0.456	0.752	0.739	1.025	1.031	1.283	1.301
1FS7_A	ASN	THR	LEU	ARG	THR	GLY	ALA	PRO	VAL	ASP	GLY	LYS	THR	GLY	PRO	LEU	PRO	SER
	1.03	1.366	1.014	1.59	1.191	0.851	0.106	0.409	-	-	-	-	0.231	0.336	0.966	1.553	1.388	0.778
									0.625	0.185	0.546	0.717						
1CGH	ASN	ASN	VAL	ALA	HIS	GLY	ILE	VAL	SER	TYR	GLY	LYS	SER	SER	GLY	VAL	PRO	PRO
_A	-	0.451	0.65	1.253	0.764	1.698	2.126	2.004	1.437	0.605	-	-	-	-	-	-	-	-
	0.141										0.456	0.473	0.482	0.649	0.037	0.698	0.306	0.516
1CYN_A	GLY	ASP	PHE	THR	ARG	GLY	ASP	GLY	THR	GLY	GLY	LYS	SER	ILE	TYR	GLY	GLU	ARG
	1.124	0.244	0.297	-	-	-	-0.73	-	-	-	-0.24	-	-	-	-	-	-	-
				0.946	0.965	0.298		0.366	0.388	0.228		0.085	0.121	0.433	0.388	1.247	0.889	0.552
1HYL_A	GLU	SER	THR	ILE	CYS	GLY	ASP	THR	SER	ASP	GLY	LYS	SER	PRO	CYS	PHE	GLY	ASP
	-	-	0.358	0.868	0.611	0.557	-	-	-	-	-	-	0.277	1.302	0.868	0.515	0.777	1.655
	0.555	0.014					0.473	0.051	1.191	0.692	1.044	0.311						
1ISA_A	ASN	ASN	LEU	ILE	LYS	GLY	THR	ALA	PHE	GLU	GLY	LYS	SER	LEU	GLU	GLU	ILE	ILE
	-	-	-	-	-	-	-	-	-	-	-	-	-	0.769	0.034	-1.04	-	0.435
	1.063	1.615	1.107	1.048	1.915	2.186	2.067	2.333	1.314	2.135	1.455	1.428	0.691				0.215	
1JAE_A	SER	GLY	GLU	LEU	SER	GLY	GLY	SER	CYS	THR	GLY	LYS	SER	VAL	THR	VAL	GLY	ASP
	-	-	0.01	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	0.524	0.566		0.872	1.187	2.068	1.835	1.582	0.925	0.865	0.954	0.912	1.488	1.039	1.874	1.196	2.024	1.857
1DDZ_A	VAL	ALA	GLN	PRO	ALA	GLY	GLN	ALA	MET	PRO	GLY	LYS	SER	ASN	ILE	PHE	ALA	ASN
	-	-	-0.87	-	-0.87	-	-	-	-	-	-	-	-	-	0.7	0.051	-	-
	0.856	1.327		0.837		0.417	0.471	0.293	0.543	0.097	0.829	0.707	0.176	0.246			0.655	0.057
1DYW	ARG	ALA	LEU	CYS	THR	GLY	GLU	ASN	GLY	ILE	GLY	LYS	SER	GLY	LYS	PRO	LEU	HIS
_A	0.51	-	0.918	1.296	-	-	-	-	-	-	-	-	-	-	-	-0.9	-	0.192
		0.153			0.229	0.951	1.062	1.374	1.634	1.748	1.454	1.428	1.032	2.096	0.941		0.229	
1QHB_A	GLN	ASN	ILE	ALA	ASP	GLY	ASP	VAL	SER	PRO	GLY	LYS	SER	PHE	LEU	LEU	PRO	MET
	-	-	-	-	-	-	-	-	-	-1.01	-	-	-	0.197	0.575	1.156	0.589	0.944
	0.036	0.579	0.903	0.938	0.061	0.838	0.485	0.855	0.801		1.099	0.647	0.342					
1CA1_A	LYS	GLU	TYR	ALA	ARG	GLY	PHE	ALA	LYS	THR	GLY	LYS	SER	ILE	TYR	TYR	SER	HIS
	1.476	0.553	1.191	1.426	0.836	0.261	0.973	0.671	-	-	-	-	-	-	-	-	-	-
									0.269	0.028	0.024	0.321	0.926	0.746	0.551	0.956	1.452	1.455
1F07 _A	PRO	LEU	ILE	LYS	GLU	GLY	ALA	GLU	ALA	ALA	GLY	LYS	SER	ILE	ALA	ASP	ILE	ASP
	-0.52	-	0.847	-	-0.49	-	-	-	-	-1.13	-	-	-0.75	0.01	-0.8	0.069	0.656	1.279
		0.171		0.494		0.405	0.339	0.939	1.171		1.507	0.448						
4PGA	ALA	TRP	GLY	PRO	LEU	GLY	MET	VAL	VAL	GLU	GLY	LYS	SER	TYR	TRP	PHE	ARG	LEU
_A	0.237	0.752	0.74	0.481	0.889	-	0.111	0.116	-0.56	-0.79	-	-	-	-	0.315	-	0.131	0.018
						0.077					0.601	0.641	0.202	0.609		0.296		
H (%)	13.64	4.55	22.73	27.27	13.64	9.09	13.64	9.09	4.55	9.09	0	0	4.55	9.09	9.09	13.64	13.64	18.18