Supporting information

Prediction of Ordered Water Molecules in Protein Binding Sites from Molecular Dynamics Simulations: The Impact of Ligand Binding on Hydration Networks

Axel Rudling[1], Adolfo Orro[1], and Jens Carlsson[2,]*

[1] Science for Life Laboratory, Department of Biochemistry and Biophysics, Stockholm University, SE-106 91 Stockholm, Sweden.

[2] Science for Life Laboratory, Department of Cell and Molecular Biology, Uppsala University, BMC, Box 596, SE-751 24 Uppsala, Sweden.

*Corresponding author e-mail: jens.carlsson@icm.uu.se.

Table of Contents

Supporting Tables

Table S1. Protonation states of histidine residues present within the simulation sphere.

Protein	Residue number	Protonation
Acetylcholinesterase	159	δ
	170	ε
	185	δ
Coagulation factor VII	25	δ
	48	ε
	115	δ
	132	δ
Fatty acid binding protein adipocyte	132	ε
Heat shock protein HSP 90-alpha	51	ε
	118	ε
	151	ε
	158	ε
Tyrosine-protein kinase JAK2	6	δ
	40	δ
	54	ε
	86	δ

	105	ε
Poly [ADP-ribose] polymerase-1	23	δ
	49	δ
	52	ε
	67	δ
	114	δ
	137	δ
	140	δ
	146	δ
Serine/threonine-protein kinase PLK1	52	ε
	57	ε
	59	δ
	64	δ
	86	ε
Protein-tyrosine phosphatase 1B	10	δ
	59	δ
	88	ε
	90	ε
	113	δ
GAR transformylase	33	ε

	76	ε
	89	δ
	103	δ
	135	ε
Muscle glycogen phosphorylase	50	ε
	55	δ
	67	δ
	78	δ
	87	ε and δ
	106	δ
	144	δ
	145	δ
Thrombin	29	δ
	52	δ
	151	δ
Trypsin I	17	ε
	31	δ
	53	ε

Table S2. Number of crystal structures of the proteins from the PDB with resolution ≤ 2.5 Å and an RMSD < 1.0 Å to the reference structure.

Protein	Number of structures
Acetylcholinesterase	56
Coagulation factor VII	53
Fatty acid binding protein adipocyte	25
Glutamate ionotropic receptor AMPA subunit 2	38
Heat shock protein HSP 90-alpha	196
Tyrosine-protein kinase JAK2	32
Poly [ADP-ribose] polymerase-1[a]	16
Serine/threonine-protein kinase PLK1[a]	8
Protein-tyrosine phosphatase 1B	116
GAR transformylase[a]	19
Muscle glycogen phosphorylase	167
Thrombin	215
Trypsin I	371
Sum:	1312

[a] This protein was not included in the analysis of frequently observed crystal waters because there were ≤ 20 crystal structures of the protein.

Table S3. Summary of results for eight apo crystal structures.

Protein	PDB code	Crystal waters[a]	Hydration sites[b]
Acetylcholinesterase	1ea5	13 (8)	19
Coagulation factor VII	1klj	1 (0)	31
Fatty acid binding protein adipocyte	3q6l	16 (14)	21
Glutamate ionotropic receptor, AMPA subunit 2	4o3b	12 (10)	25
Heat shock protein 90-alpha	1uyl	15 (13)	17
Protein-tyrosine phosphatase 1B	2cm2	12 (10)	19
Thrombin	2uuf	20 (14)	29
Trypsin I	1s0q	19 (9)	26
Sum:	-	108 (78)	187

[a] Number of waters in the binding site of the crystal structure, *i.e.* within 4 Å of both the protein and ligand (from the reference structure). The number of crystal waters that were reproduced by hydration sites is shown in parentheses.
[b] Number of hydration sites within 4 Å of the protein and ligand (from the reference structure).