Original Research Article **Bioinformatics**

International Journal of Pharma and Bio Sciences

ISSN 0975-6299

STRUCTURAL DISCRIMINATION OF PURINES AND PYRIMIDINES BY PROTEINS THROUGH WATER-MEDIATED CONTACTS

S. USHA*[1] AND KM. SARAVANAN[2]

[1] Department of Bioinformatics, School of Life Sciences, Bharathidasan University, Tiruchirappalli, Tamilnadu- 620 024, India.
[2]Centre of Excellence in Bioinformatics, School of Biotechnology, Madurai Kamaraj University, Madurai - 625021.

ABSTRACT

Water molecules play an important role in the recognition and stabilization of protein-ligand interactions. The water molecules at binding sites bridge the protein and ligand and can make three or more hydrogen bonds when distance and bond angles are used as criteria to define hydrogen-bonding interactions. There exists a distinct structural similarity between purines and between pyrimidine moieties. In the present study, we have analyzed the contribution of the water-mediated hydrogen bonds in the structure wise discrimination of these nucleotide bases. Significant differences in the amino acid residue preferences and the ligand atom preferences were observed in the case of purine moieties, adenine and guanine. In pyrimidine nucleotides, the amino acid residue Serine was preferred in both cytosine and thymine contacts; Asparagine was preferred in both thymine and uracil; Aspartic acid was preferred in both uracil and cytosine water-mediated contacts. Cytosine O_2 atom was highly preferred whereas the O_4 atom was highly preferred in thymine and uracil interactions. Since, the water networks provide increased specificity and affinity, the ability to include water-binding sites into the interface between a drug and its target might prove useful in drug design strategies.

KEYWORDS: **Water-mediated, purine nucleotides, pyrimidine bases, protein-ligand interactions, structural discrimination**

S. USHA
Department of Bioinformatics, School of Life Sciences, Bharathidasan University, Tiruchirappalli, Tamilnadu- 620 024, India.

* **Corresponding author**

INTRODUCTION

Water molecules present at the interface of protein-ligand complexes contribute to their binding affinity and specificity[1]. Water molecules can stabilize the protein-ligand complexes by mediating hydrogen bond interactions between the protein and the ligand[2-5] or by being displaced upon ligand binding[6,7]. Earlier studies have also reported that inclusion of bridging water molecules was also found to improve the binding affinity[8] as well as docking predictions[9]. Hence, the significant role of interfacial water molecules should be considered important in protein-ligand binding[10-12].

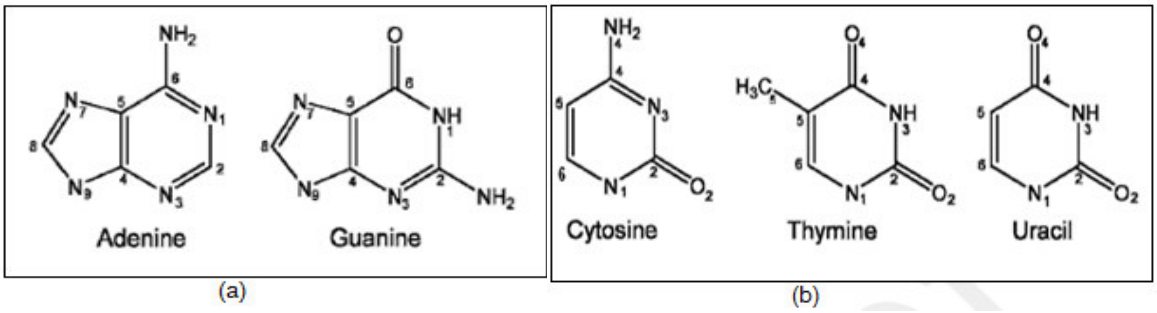

Figure 1
(a) Structures of adenine and guanine bases (b) Structures of cytosine, thymine and uracil bases

The purine bases adenine and guanine have similar arrangement of atoms in their structures except in their first, second, and sixth positions and the pyrimidine derivative molecules, viz. cytosine, thymine, and uracil, resemble each other in their structures except the atoms in the first, second, and sixth atom positions (Figure 1). In our previous publications, we have analysed the structure wise discrimination of pyrimidine[13] (Usha and Selvaraj, 2014) and purine[14] (Usha and Selvaraj, 2015) basal structures in terms of their non-bonded interactions. In continuation with our previous works, in the present study, we have carried out such discriminative analysis based on the contribution made by the interfacial water molecules[15]. The role of water molecules that mediate hydrogen bonds between proteins and ligands were analyzed to know their contribution to the structural discrimination of the similar ligand moieties, adenine and guanine as well as cytosine, thymine and uracil.

MATERIALS AND METHODS

Water-mediated contacts were obtained from the PEARLS (Program for Energetic Analysis of Receptor-Ligand System) server[16] (http://ang.cz3.nus.edu.sg/cgi-bin/prog/rune.pl) for the purine-protein and pyrimidine-protein complexes. The contribution of ligand-water-receptor hydrogen bonding in PEARLS has been estimated by computing the energy of each of the ligand-water and water-receptor hydrogen bond using the Morse potential[17]. The dataset of the current analysis was carried out using 145 adenine-protein, 128 guanine-protein, 53 cytosine-protein, 43 thymine-protein and 115 uracil-protein complexes. The list of PDB entries included for the dataset the purine- and pyrimidine-protein complexes for water-mediated interaction analysis are provided in the Appendix.

RESULTS

We have carried out an analysis of water-mediated protein-ligand interactions to understand the role of water molecules in the structural discrimination between purines and among pyrimidines by proteins.

Water-mediated interactions
In purine-protein complexes

There were 93 water-mediated contacts in adenine-protein complexes and 126 contacts in guanine-protein complexes. The amino acid residues Threonine (12.9%), Arginine (9.7%) and Valine (8.6%) are highly preferred in the water mediated contacts of adenine. Glycine (13.5%), Aspartic acid (11.1%) and Threonine (9.5%) are preferred in guanine water-mediated contacts (Figure 2).

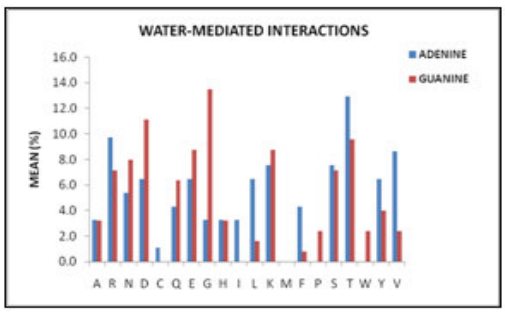

Figure 2
Amino acid residue contacts with purine bases through water-mediated interactions
Water-mediated hydrogen bond energies of the complexes were found to range from -0.12 to -0.22 Kcal/mol and their mean bond energy was -0.19 Kcal/mol for both adenine- and guanine-protein complexes.

Table 1
Precedence of purine atoms to form water-mediated contacts with proteins

Adenine atoms	Precedence (%)	Guanine atoms	Precedence (%)
N_7	34.4	O_6	29.4
N_6	25.8	N_2	24.6
N_1	23.7	N_7	19.0
N_3	15.1	N_3	13.5
N_1	0.0	N_1	12.7
N_9	1.1	N_9	0.8

N_7 atoms (34.4%) form predominant contacts in adenine water mediated contacts, followed by N_6 and N_1 atoms (25.8% and 23.7% respectively) of adenine whereas guanine O_6 (29.4%) and N_2 (24.6%) atoms highly participate in the water mediated interactions of guanine-protein complexes (Table 1).

In pyrimidine-protein complexes

38 water-mediated contacts in cytosine-protein complexes, 42 contacts in thymine-protein complexes and 78 contacts in uracil-protein complexes were observed.

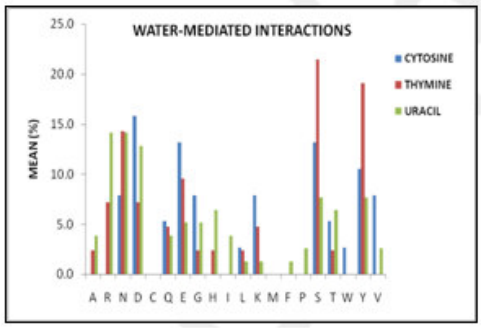

Figure 3
Amino acid residue contacts with pyrimidine bases through water-mediated interactions

Aspartic acid (15.8%), Glutamic acid (13.2%) and Serine (13.2%) were preferred in cytosine water mediated contacts. Thymine water mediated contacts are favoured by Serine (21.4%), Threonine (19.0%) and Asparagine (14.3%) residues. Arginine and Asparagine (14.1%), and Aspartic acid (12.8%) prefer water-mediated interactions in uracil-protein complexes (Figure 3). Energy of water-mediated hydrogen bonds were found to range between -0.14 and -0.22 Kcal/mol for cytosine- and thymine-protein complexes and -0.13 to -0.22 Kcal/mol for uracil-protein complexes. The mean water-mediated hydrogen bond energies were found to be -0.19 Kcal/mol for cytosine- and uracil-protein complexes and -0.20 Kcal/mol for thymine-protein complexes.

Table 2
Precedence of pyrimidine atoms to form water-mediated contacts with proteins

Cytosine atoms	Precedence (%)	Thymine atoms	Precedence (%)	Uracil atoms	Precedence (%)
O_2	50.0	O_4	42.9	O_4	56.4
N_4	31.6	O_2	33.3	O_2	26.9
N_3	18.4	N_3	21.4	N_3	14.1
N_1	0.0	N_1	2.4	N_1	2.6

The O_2 (50.0%) atom of cytosine contribute to half of the water mediated interactions of cytosine; followed by N_4 atom (31.6%). Whereas the precedence of atoms to form water mediated contacts are found to be similar for thymine and uracil i.e., O_4 (Thymine: 42.9% and Uracil: 56.4%) and O_2 (Thymine: 33.3% and Uracil: 26.9%) respectively (Table 2).

DISCUSSION

Water molecules conserved in the ligand-bound structures generally participate in water-mediated hydrogen bonds between the protein and the ligand and contribute for the recognition of ligands by proteins[10]. Such a hydrogen-bonded network of water molecules causes stability of the proteins to be complexed with only the cognate ligand, thus contributing to the specificity of ligand recognition[18-19].The importance of water-mediated interactions in ligand recognition by proteins has been reported in many earlier studies[20-24]. The water molecules have also shown their impact in many diseases associated systems. For example, a study on the HIV-I virus has demonstrated that the direct interaction of a proline residue of Gag polyprotein with the hydrophobic pocket of cyclophilin ligand is essential for the formation of the Gag-cyclophilin complex, which is a crucial step in the HIV-1 virus life cycle[25]. In another case, the hydrophobic pocket of the glycoprotein dengue virus type 2 binds to a hydrophobic ligand for membrane fusion, thus providing opportunities for designing dengue virus inhibitors against the hydrophobic pocket of the virus protein[26]. Lu *et al*. (2007)[27] have analyzed the ligand-bound water molecules in 392 protein-ligand complexes. They have shown that the amino acids with charged side chains, glutamic acid, arginine, aspartic acid, and lysine exhibit highest hydration propensities at the protein-ligand interfaces. This shows that the hydration propensities

depend upon the nature of atom charges, atoms with positive or negative charges have high hydration propensity whereas it is low for neutral and hydrophobic atoms. Accordingly, the water-mediated interactions in purine- and pyrimidine-protein complexes are dominated by amino acids containing charged groups; arginine in adenine-, aspartic acid in guanine-; aspartic acid and glutamic acid in cytosine-; asparagine in thymine- and arginine, asparagine and aspartic acid in uracil-water-protein contacts. Also, the amino acid preferences as well as the ligand atom preferences to form water-mediated contacts are found to vary between similar ligand moieties (between purine bases as well as among pyrimidine bases) (Figures 2 and 3). Several docking studies have shown that the inclusion of all crystallographic water molecules with the ligand atoms resulted in improved accuracy of the docking predictions[28-30]. The empirical scoring function HINT12 was used to estimate the free energy of binding in HIV-1 protease-ligand complexes and the correlation between HINT scores and the experimentally determined binding constants was found to significantly improve when an important interface bridging water molecule was included. Knowing the contribution of the important water molecular interactions with proteins, solvated energy minimized structure was used for Molecular Dynamics simulation of Inosine 5'- Monophosphate dehydrogenase-II (human) structure[31]. These studies suggest that the interfacial water molecules play a crucial role in protein-ligand binding and should be considered during drug design processes.

CONCLUSION

Water-mediated interactions analysis reveals the structural discrimination between purine and among pyrimidine ligands in terms of amino acid and ligand atom preferences. Protein-ligand recognition phenomenon could be best understood on the basis of a proper description of water molecules. Hence, development of new implicit solvent models[32] (Setny et al., 2009) shall provide a realistic representation of solvation properties. Water molecules constitute an important component in structure-based drug design and ligand binding affinity predictions could be improved by mimicking, displacing, and targeting bound water molecules[33] (Marrone, Briggs and McCammon, 1997). The present analysis provides the evidence of water molecules in the discrimination of the structurally similar nucleobases.

ACKNOWLEDGEMENT

We thank all the crystallographers who contributed valuable data to Protein Data Bank

CONFLICT OF INTEREST

Conflict of interest declared none.

REFERENCES

1. Ladbury JE. Just add water! The effect of water on the specificity of protein-ligand binding sites and its potential application to drug design. Chem Biol. 1996 Dec 31;3(12):973-80.
2. Bhat TN, Bentley GA, Boulot G, Greene MI, Tello DW, Dall'Acqua W, Souchon H, Schwarz FP, Mariuzza RA, Poljak RJ. Bound water molecules and conformational stabilization help mediate an antigen-antibody association. Proc Natl Acad Sci U S A. 1994 Feb 1;91(3):1089-93.
3. Varghese JN, Chandana Epa V, Colman PM. Three-dimensional structure of the complex of 4-guanidino-Neu5Ac2en and influenza virus neuraminidase. Protein Sci. 1995 Jun 1;4(6):1081-7.
4. Meiering EM, Wagner G. Detection of long-lived bound water molecules in complexes of human dihydrofolate reductase with methotrexate and NADPH. J Mol Biol. 1995 Mar 24;247(2):294-308.
5. Smith BJ, Colman PM, Von Itzstein M, Danylec B, Varghese JN. Analysis of inhibitor binding in influenza virus neuraminidase. Protein Sci. 2001 Apr 1;10(4):689-96.
6. Lam PY, Jadhav PK, Eyermann CJ, Hodge CN, Ru Y, Bacheler LT, Meek JL, Otto MJ, Rayner MM, Wong YN. Rational design of potent, bioavailable, nonpeptide cyclic ureas as HIV protease inhibitors. Science. 1994 Jan 21;263(5145):380-4.
7. De Lucca GV, Erickson-Viitanen S, Lam PY. Cyclic HIV protease inhibitors capable of displacing the active site structural water molecule. Pharm Biotechnol. 1997 Jan 31;2(1):6-18.
8. Fornabaio M, Spyrakis F, Mozzarelli A, Cozzini P, Abraham DJ, Kellogg GE. Simple, intuitive calculations of free energy of binding for protein-ligand complexes. 3. The free energy contribution of structural water molecules in HIV-1 protease complexes. J Med Chem. 2004 Aug 26;47(18):4507-16.
9. Rarey M, Kramer B, Lengauer T. The particle concept: placing discrete water molecules during protein-ligand docking predictions. Proteins. 1999 Jan 1;34(1):17-28.
10. Poornima CS, Dean PM. Hydration in drug design. 3. Conserved water molecules at the ligand-binding sites of homologous proteins. J Comput Aided Mol Des. 1995 Dec 1;9(6):521-31.
11. Hamelberg D, McCammon JA. Standard free energy of releasing a localized water molecule from the binding pockets of proteins: double-decoupling method. J Am Chem Soc. 2004 Jun 23;126(24):7683-9.
12. Lu Y, Yang CY, Wang S. Binding free energy contributions of interfacial waters in HIV-1 protease/inhibitor complexes. J Am Chem Soc. 2006 Sep 13;128(36):11830-9.
13. Usha S, Selvaraj S. Structure-wise discrimination of cytosine, thymine, and uracil by proteins in terms of their nonbonded interactions. J Biomol Struct Dyn. 2014 Oct 3;32(10):1686-704.

14. Usha S, Selvaraj S. Structure-wise discrimination of adenine and guanine by proteins on the basis of their nonbonded interactions. J Biomol Struct Dyn. 2015 Jul 3;33(7):1474-92.
15. Usha S. Role of water-mediated interactions and solvent accessible surface area in the structural discrimination between purines and pyrimidines by proteins [Thesis (Chapter IV]. Bharathidasan University; 2014.
16. Han LY, Lin HH, Li ZR, Zheng CJ, Cao ZW, Xie B, Chen YZ. PEARLS: program for energetic analysis of receptor-ligand system. J Chem Inf Model. 2006 Jan 23;46(1):445-50.
17. Baird NC. Simulation of hydrogen bonding in biological systems: Ab initio calculations for NH3-NH3 and NH3-NH4+. Int J Quantum Chem. 1974 Jan 17;8(S1):49-54.
18. Rejto PA, Verkhivker GM. Mean field analysis of FKBP12 complexes with FK506 and rapamycin: Implications for a role of crystallographic water molecules in molecular recognition and specificity. Proteins. 1997 Jul 1;28(3):313-24.
19. Wester MR, Johnson EF, Marques-Soares C, Dansette PM, Mansuy D, Stout CD. Structure of a substrate complex of mammalian cytochrome P450 2C5 at 2.3 Å resolution: evidence for multiple substrate binding modes. Biochemistry. 2003 Jun 3;42(21):6370-9.
20. Uchida T, Ishimori K, Morishima I. The effects of heme pocket hydrophobicity on the ligand binding dynamics in myoglobin as studied with leucine 29 mutants. J Biol Chem. 1997 Nov 28;272(48):30108-14.
21. Carey C, Cheng YK, Rossky PJ. Hydration structure of the α-chymotrypsin substrate binding pocket: the impact of constrained geometry. Chem Phys. 2000 Aug 15;258(2):415-25.
22. Levy Y, Onuchic JN. Water mediation in protein folding and molecular recognition. Annu Rev Biophys Biomol Struct. 2006 Jun 9;35:389-415.
23. Young T, Abel R, Kim B, Berne BJ, Friesner RA. Motifs for molecular recognition exploiting hydrophobic enclosure in protein–ligand binding. Proc Natl Acad Sci U S A. 2007 Jan 16;104(3):808-13.
24. Qvist J, Davidovic M, Hamelberg D, Halle B. A dry ligand-binding cavity in a solvated protein. Proc Natl Acad Sci U S A. 2008 Apr 29;105(17):6296-301.
25. Braaten D, Ansari H, Luban J. The hydrophobic pocket of cyclophilin is the binding site for the human immunodeficiency virus type 1 Gag polyprotein. J Virol. 1997 Mar 1;71(3):2107-13.
26. Modis Y, Ogata S, Clements D, Harrison SC. A ligand-binding pocket in the dengue virus envelope glycoprotein. Proc Natl Acad Sci U S A. 2003 Jun 10;100(12):6986-91.
27. Lu Y, Wang R, Yang CY, Wang S. Analysis of ligand-bound water molecules in high-resolution crystal structures of protein-ligand complexes. J Chem Inf Model. 2007 Mar 26;47(2):668-75.
28. Hartshorn MJ, Verdonk ML, Chessari G, Brewerton SC, Mooij WT, Mortenson PN, Murray CW. Diverse, high-quality test set for the validation of protein-ligand docking performance. J Med Chem. 2007 Feb 22;50(4):726-41.
29. Roberts BC, Mancera RL. Ligand-protein docking with water molecules. J Chem Inf Model. 2008 Feb 25;48(2):397-408.
30. Thilagavathi R, Mancera RL. Ligand− protein cross-docking with water molecules. J Chem Inf Model. 2010 Feb 17;50(3):415-21.
31. Mishra DK, Mukhopadhyay BP, Bairagya HR. Molecular modeling of inosine 5'-monophosphate dehydrogenase-ii (human) structure using MDSIMULATION METHOD. Int J Pharm Bio Sci. 2012;3:102-20.
32. Setny P, Wang Z, Cheng LT, Li B, McCammon JA, Dzubiella J. Dewetting-controlled binding of ligands to hydrophobic pockets. Phys Rev Lett. 2009 Oct 30;103(18):187801.
33. Marrone TJ, Briggs, and JM, McCammon JA. Structure-based drug design: computational advances. Annu Rev Pharmacol Toxicol. 1997 Apr;37(1):71-90.