A Comprehensive Analysis of Protein Data Bank reveals Low Desolvation Penalty in π-Cation System

Jing-Fang Yang
Central China Normal University https://orcid.org/0000-0002-3507-1664
Meng-Yao Wang
Central China Normal University https://orcid.org/0000-0003-2664-9214
Di Wang
Central China Normal University https://orcid.org/0000-0003-0870-3934
Jing-Yi Li
Central China Normal University https://orcid.org/0000-0002-5672-5209
Ge-Fei Hao (gefei_hao@foxmail.com)
Central China Normal University
Qing X. Li
University of Hawai'i at Manoa Libraries https://orcid.org/0000-0003-4589-2869
Guang-Fu Yang
Huazhong University of Science and Technology https://orcid.org/0000-0003-4384-2593

Research article

Keywords: desolvation penalty, π-cation, cation-π, molecular interaction, drug and pesticide design

Posted Date: August 28th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-62381/v1

A comprehensive analysis of Protein Data Bank reveals low desolvation penalty in π-cation system

Jing-Fang Yang [b,c,†], Meng-Yao Wang [b,c,†], Di Wang[b,c], Jing-Yi Li[b,c], Ge-Fei Hao[a,*], Qing X. Li[d,*], and Guang-Fu Yang[b,c,e]

[a]*State Key Laboratory Breeding Base of Green Pesticide and Agricultural Bioengineering, Key Laboratory of Green Pesticide and Agricultural Bioengineering, Ministry of Education, Research and Development Center for Fine Chemicals, Guizhou University, Guiyang 550025, P. R. China;* [b]*Key Laboratory of Pesticide & Chemical Biology, Ministry of Education, College of Chemistry, Central China Normal University, Wuhan 430079, P.R.China;* [c]*International Joint Research Center for Intelligent Biosensor Technology and Health, Central China Normal University, Wuhan, 430079, China;* [d] *Department of Molecular Biosciences and Bioengineering, University of Hawaii at Manoa, Honolulu, Hawaii 96822, United States;* [e]*Collaborative Innovation Center of Chemical Science and Engineering, Tianjin 300072, P.R.China*

Correspondence:

Qing X. Li, Ph.D.

Professor

Department of Molecular Biosciences and Bioengineering, University of Hawaii at Manoa, Honolulu, Hawaii 96822, United States

E-mail: qingl@hawaii.edu

[†]Co-first authors.

*To whom correspondence should be addressed. *E-mail: gefei_hao@foxmail.com; qingl@hawaii.edu*

Abstract

Cation-π interactions widely exist between ligand-protein interfaces, attracting much attention in molecular recognition in recent years. Interactions named cation-π and π-cation (cationic vs arene small molecular ligands) shall be separately considered in drug and pesticide design process. The two interactions involved in ligands and protein pockets may differ in energy features and therefore offers significant inspiration for drug and pesticide design. However, an in-depth study on differences between cation-π and π-cation systems from an energy perspective is still lacking. In this study, we calculated and compared cation-π and π-cation systems in terms of physicochemical properties of ligand/protein and solvation effect. It seems that the desolvation penalty of the cation-π systems was relatively higher than the π-cation pairs, even though these interactions both can improve the ligand activity. This is the reason for evolution converged on π-cation interactions in the cation-π-mediated proteins. The π-cation interaction facilitating the inhalation of ligand to the pocket may provide a new sight for the molecular design of pharmaceuticals and pesticides.

Key Words: desolvation penalty; π-cation; cation-π; molecular interaction; drug and pesticide design

Introduction

The cation-π interaction, dominated by the electrostatic attraction between electron-rich aromatic rings and positively charged groups, is an essential topic in structure-based drug and pesticide design [1-6]. Some instances have been reported in the modulation of receptor and ligand interactions. Douglas *et al*. found that cation-π interactions can modulate the N-methyl-d-aspartate (NMDA) receptor inhibitory potencies of inhaled drugs of abuse [7]. Borissow *et al*. illustrated that the cation-π interaction between adenophostin and Arg504 of Ins(1,4,5)P3R receptor is responsible for the high potency [8]. Zhu *et al*. revealed that the cation-π interaction between carboxamide fungicides and C_R46 in the Q-site of succinate-ubiquinone oxidoreductase contributes significantly to affinity [9]. Therefore, the cation-π interaction plays an important role in modulating molecular recognition between protein and ligand.

The definition of the interactions between cation and π in biological systems has experienced a long and complex process. Based on the initial definition from Dougherty *et al*., cation-π interactions can arise between ligands and tryptophan (Trp), tyrosine (Tyr), phenylalanine (Phe), arginine (Arg), and lysine (Lys) [6, 10-12]. Being either cation or π ligand can classify such interactions into cation-π and π-cation pairs in drug or pesticide and target interactions (Fig. 1) [13]. The π element in a cation-π pair is typically provided by the side-chains of aromatic residues, Trp, Tyr, and Phe, while the positively charged ligands act as “cation”. For π-cation pairs, two protonated amino acids (Arg and Lys) are “cation”, with arene ligands providing the π element [14].

Recent attention has been paid to study various weak interactions in computational based drug and pesticide design [15-17]. For example, Wendler *et al*. put forward fitting functions by using different parameters to evaluate energies and detect hydrogen bonds [18]. The fitting functions describe hydrogen bonds in amino acid dimers and the cooperativity of hydrogen bond energies in water clusters quite well [18]. Chourasia *et al*. constructed an *Aromatic-Aromatic Interactions Database* and analyzed the connectivity patterns of π–π networks in proteins [19]. A cluster arrangement of aromatic residues represents a stronger propensity than a linear arrangement. Du *et al*. proposed an empirical formulation for the interaction energy calculation of cation-π in proteins, which can quantitatively determine cation-π interactions [20]. The computational approach can be precise in drug and pesticide design. However, the energetic difference between cation-π and π-cation in drug and protein interaction systems remains unclear. While understanding the difference between the two systems makes it easier to use cation-π and π-cation systems for rational drug and pesticide design. Therefore, there is a strong need to perform a systematic analysis of the energetic difference between these two interactions in all available protein structures.

In this study, we get 1334 cation-π systems and 2174 π-cation systems from 141,706 crystal structures in the PDB database by using Protein-Ligand Interaction Profiler (PLIP) [21]. The physicochemical properties and desolvation penalty were then calculated. Firstly, the physicochemical properties, such as hydrophobicity and the negative decadic logarithm of the ionization constant (pKa) of proteins and ligands, were studied [22-24]. Then, the binding energy was compared. The results showed that π-cation pairing can sharply reduce the energy of the desolvation penalty, which can facilitate the inhalation of the drug or pesticide into the pocket

and promote the coevolution of proteins and ligands. The reduced desolvation penalty may be induced by the physiochemical property difference between π-supplier and cation-supplier. Therefore, taking a proper π-supplier of the ligand scaffold into account will guide how to design potent drugs and pesticides.

Methods

The definition of cation-π and π-cation interactions

The PLIP was used to geometrically identify cation-π interactions with default parameters [21]. Furthermore, the cation-π and π-cation interactions were classified by the ligand groups. Tertamine, quartamine, guanidine, and sulfonium ligands provided cations for cation-π interactions. Aromatic ligands were π components in π-cation interactions.

The physicochemical property evaluation

The hydrophobicity of ligand binding pockets was obtained with Fpocket [25]. The pKa was calculated using Hammett-Taft equation derived from plenty of libraries of experimental data. The graphs of descriptors of proteins and ligands related to cation-π and π-cation interactions were analyzed and edited by Origin2019b [26, 27].

Structural optimization

The Sander module of Amber16 was used to minimize the energy of complex structures with implicit solvent [28]. The AMBER ff14SB force field was used for residues, and the general

AMBER force field (gaff) was used for ligands [29, 30]. The cutoff distance for the long-range electrostatic interaction was set at 10.0Å. The minimization procedure consisted of the following three steps [31]. First, only hydrogens, ions, and water molecules were allowed to move, and the solute remained fixed with a constraint of 500 kcal mol^{-1} $Å^{-2}$. Then, the backbone atoms of the protein were fixed, and other atoms were relaxed. Finally, all the atoms of the system were free to move. In each step, the steepest descent method was used for the first 2000 cycles and the conjugate gradient method was used for the next 1000 cycles to perform energy minimization [31].

Desolvation penalty evaluation

The desolvation penalty (SOL_value) analysis was performed based on Molecular Mechanics Poisson-Boltzmann Surface Area (MM/PBSA), a common method for binding free energy calculation, which is performed using Amber16 software [32].

The protein-ligand complex binding energy (ΔG_{bind}) was estimated by the molecular mechanical (MM) gas-phase binding energy (ΔE_{MM}) and solvation energy (desolvation penalty) (ΔG_{solv}) shown in Equation 1 [27].

$$\Delta G_{bind} = \Delta E_{MM} + \Delta G_{solv} \quad (1)$$

The ΔE_{MM} can be evaluated as the sum of the electrostatic energy (ΔE_{ele}), van der Waals interaction energy (ΔE_{vdw}) and the bond, angle, dihedral energies (ΔE_{int}) according to Equation 2. The ΔG_{solv} can be divided into two parts – the electrostatic desolvation penalty ($\Delta G_{PB/GB}$) and the nonpolar desolvation penalty (ΔG_{np}) (Equation 3).

$$\Delta E_{MM} = \Delta E_{int+} \Delta E_{ele} + \Delta E_{vdw} \quad (2)$$

$$\Delta G_{solv} = \Delta G_{PB/GB} + \Delta G_{np} \quad (3)$$

The SOL_value was defined as a partition ratio of the solvation energy (ΔG_{solv}) in the sum of the molecular mechanical gas-phase binding energy (ΔE_{MM}) and the solvation energy (ΔG_{solv}) by using Equation 4.

$$\text{SOL\_value} = (|\Delta G_{solv}|)/ (|\Delta E_{MM}| + |\Delta G_{solv}|) \quad (4)$$

Results and discussion

Data content and analysis

To identify the difference between cation-π and π-cation interactions, we developed a workflow to collect cation-π and π-cation complex crystal structures from Protein Data Bank (PDB) (Fig. 2a). Approximately 60000 pdb files with small molecules were filtered out from 141706 crystal structures downloaded from the PDB database. Then, 8263 complex crystal structures with cation-π and π-cation interactions were geometrically identified through the PLIP. Finally, 1334 complexes with 1926 receptor-ligand cation-π interactions and 2174 complexes involved in 2643 receptor-ligand π-cation interactions were collected and classified according to the ligand groups. The number of complex crystal structures with π-cation interactions was about 1.63 times of that with cation-π, indicating that the π-cation interaction was the more important component in the ligand and receptor interactions.

Further, the data analysis for cation-π and π-cation residue pairs was performed. The abundance

of Phe (324, 16.82%) and Tyr (327, 16.98%) was absolutely lower than that of Trp (1275, 66.20%) in cation-π interactions. For the π-cation interactions, the percentage of Arg was about 67.42%, which was relatively higher than that of Lys (32.58%, Fig. 2b). These results implicated that Trp and Arg were essential in the cation-π and π-cation interactions, which was consistent with the previous results [33-35].

Properties of protein pockets and ligands

As we all known, the chemistry community recognized the cation-π interaction as a major force for molecular recognition, joining the hydrophobic effect [36]. Therefore, the hydrophobicity, hydrophily, and other properties of proteins and ligands were evaluated.

The hydrophobicity score of ligand binding pockets was calculated based on the method proposed by Monera *et al* [22, 37]. The higher score represents higher hydrophobicity. The calculated hydrophobic scores of two systems both ranged from -15 to 65 with the normal distribution (Fig. 3a). However, there was a slight difference between cation-π and π-cation: the hydrophobic scores of π-cation pairs clustered in the interval of -5 to 45, while the most scores of cation-π moved to the interval of 5 to 45. From violin plots, we found that most score (cation-π: 79.60%, π-cation: 77.60%) lied in the range of 5-35 (Fig. 3b). Nevertheless, the shape of π-cation moved down compared with that of cation-π, revealing that the binding pockets of cation-π were much more hydrophobic.

178 The hydrogen ion is the common proton in the ligands. The pKa is always used to describe the
179 ability of acid dissociation, which is related to the solubility of the ligand [24, 38]. The smaller
180 the value of pKa indicates the stronger the acid, in turn, the stronger base[39]. Hence, the pKa of
181 ligands was studied to make the difference between cation-π and π-cation systems. Just like the
182 hydrophobic score, the ligands in π-cation pairs were the lower pKa biased (Fig. 4a). When the
183 pKa was in the range 0 to 25, the possession percentage of π-cation systems was lower than the
184 cation-π. For the range -15 to 0, the possession percentage of π-cation systems was much larger
185 than that of cation-π. Fig. 4b showed the different shapes of two violin plots. In the π-cation
186 system, there were two wide shapes in the interval of -10 to 0 and the interval of 0 to 10. For the
187 cation-π systems, three wider sections were shown. Meanwhile, the first wide part of cation-π
188 was relatively higher than that of π-cation. On the other hand, the last wide shape of π-cation was
189 lower than the cation-π. The median of cation-π (~10) was higher than π-cation (~5). These
190 results indicated that the basic ligands preferred to act as cations in cation-π systems.
191

192 **Desolvation penalty difference**

193 Solvation energy is a fundamental thermodynamic quantity to estimate the desolvation cost of a
194 ligand-binding with a protein [40]. The effect of the desolvation penalty is important in drug
195 discovery due to its influence on the inhalation of a drug into the pocket [41]. Accounting for the
196 effect of solvent on the strength of molecular interactions has been a long-standing problem for
197 structure-based drug design. From the above, we found that the strong basic ligands preferred to
198 bind with the hydrophobic binding pockets to form cation-π, vice versa, the π-cation interaction
199 would be formed. Hence, the solvent may be a dominant force to dictate the interaction of
200 molecules and proteins with the opposite solubilities. Actually, the solvent exposure phenomenon

is frequently found in the residues involved in cation-π or π-cation interactions, and the surrounding solvents may be one of the main factors to modulate the strength of cation-π or π-cation interactions [42, 43]. In particular, some theoretical studies have established the importance of the desolvation penalty for cation-π systems [42, 44-47]. However, whether there is a difference in the desolvation penalty between cation-π and π-cation systems is still unknown.

In this study, the SOL_value, avoiding the system interference, was defined to determine the overall influence of the desolvation penalty for the binding affinity according to Equation 4. If the SOL_value is close to 0.50, the absolute value of ΔG_{solv} is clearly close to the absolute value of ΔE_{MM}. In contrast, if the SOL_value tends to 0.00, the ΔG_{solv} is definitely close to 0.00, resulting in a lower desolvation penalty for the binding affinity.

The distributions of SOL_value of 1334 complexes with cation-π interactions and 2174 complexes with π-cation interactions were analyzed. As shown in Fig. 5a, the SOL_value ranged from 0.00 to 0.50 with 0.45 as a demarcation point. When the SOL_value was larger than 0.45, there was more cation-π interaction (45.50%) than π-cation interaction (27.37%). When the SOL_value is ranged from 0.45 to 0.47, the percentage of cation-π (26.76%) was much larger than that of π-cation (13.01%). In contrast, the percentage of π-cation was larger than the cation-π in the SOL_value ranges of 0.00~0.30, 0.30~0.40, and 0.40~0.45. The percentages of π-cation were 25.71%, 22.63%, and 24.29%, respectively. The largest difference appears in the SOL_value range of 0.00~0.30, in which the percentage of π-cation was 10.64% higher than that of cation-π. The bottom of the violin plot of cation-π was much thinner than that of π-cation systems, which demonstrates that there were fewer values in comparison to π-cation systems

when tending to 0.00 (Fig. 5b). In summary, the SOL_values of over 72% of π-cation systems are much closer to 0.00. This means that the π-cation interactions in the complexes may result in a lower desolvation penalty which is conducive to the binding affinity. It is maybe the reason for the enrichment of π-cation in the natural systems.

Case study

The above results revealed that the hydrophobicity of cation-π pockets was relatively higher, while the strong basic cation ligands in cation-π pairs had a stronger ability to dissolve in the water. These properties led to the higher solvent effect in cation-π interactions. To further illustrate the values of π-cation and cation-π interactions, we collected a series of representative crystal structures with available experimental binding affinities, such as, the half-maximum inhibition concentration (IC_{50}) and the inhibitory constant (K_i) (Table S1 and S2). Further, we used two cases to explain the difference between cation-π and π-cation in the protein systems (Fig. 6a and b). Serine/threonine kinase acts as an essential component of the mitogen-activated protein kinase (MAPK) signal transduction pathway. MAPK1/ERK2 is one of the two MAPKs playing an important role in the MAPK/ERK cascade and is a key oncogenic pathway implicated in a variety of human cancers. The K_i of ligand 33A was reduced from 2300nM to 86 nM after the chlorine and benzene ring substitution [48]. The benzene ring of 33A formed π-cation interactions with Lys52. On the other hand, the leukotriene 4 hydrolase (LTA4H) is a key target for the treatment of cardiovascular disease. The binding affinity of 27P (IC_{50}=26nM) was improved after the modification of 24P (IC_{50}=87nM) [49]. The nitrogen atom formed cation-π interaction with Tyr267 and Tyr378. It was common that the hydrophobicity score of the MAPK1 pocket for π-cation pair was lower than that of LTA4H for cation-π. Meanwhile, the

SOL_value was improved after the introduction of cation-π and π-cation interactions. However, these interactions both can strengthen the interaction of the ligand with receptors, even though the SOL_value of 27P-LTA4H (0.41) was much higher than 33A-MAPK1 (0.18). Meanwhile, it should be noted that the activity fold changed of 33A (26.74) was much higher than 27P (3.35). These results are consistent with our conclusion that the π-cation interaction has a lower desolvation penalty than cation-π systems, which would contribute to the ligand binding.

Conclusion

The cation and π interactions play an essential role in drug(pesticide)-target interaction. Differentiating cation-π and π-cation systems can facilitate the rational molecular design of pharmaceuticals and pesticides. We comprehensively compared the physicochemical properties of protein pockets and ligands and solvation energy for cation-π and π-cation systems in this study. Compared with a cation-π system, a π-cation system probably results in a lower desolvation penalty, which may facilitate the inhalation of ligand to the pocket and determine the coevolution of proteins and ligands. Due to the dependence upon the binding site of the target proteins in drug and pesticide design, the π-cation interaction is a valuable tool when the ligand contains an aromatic functional group. Therefore, taking a proper π-supplier of the ligand scaffold into account can guide drug and pesticide design.

Abbreviations

NMDA: N-methyl-d-aspartate; Trp: tryptophan; Tyr: tyrosine; Phe: phenylalanine; Arg: arginine; Lys: lysine; pKa: the ionization constant; MM/PBSA: Molecular Mechanics Poisson-Boltzmann

Surface Area; PDB: Protein Data Bank; IC_{50}: the half-maximum inhibition concentration; K_i: the inhibitory constant; MAPK: the mitogen-activated protein kinase; LTA4H: the leukotriene 4 hydrolase

Acknowledgements

None.

Author's contributions

All authors have contributed to the manuscript and given approval to the final version of the manuscript.

Funding

This work was supported in part by the National Key R&D Program (2017YFD0200501) and the National Natural Science Foundation of China (21772059, 91853127, and 31960548). Program of Introducing Talents of Discipline to Universities of China (111 Program, D20023). The USDA (Hatch HAW5032-R), and Hawaii Community Foundation (18ADV-90801).

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Mecozzi S, West AP, Jr., Dougherty DA (1996) Cation-pi interactions in aromatics of biological and medicinal interest: electrostatic potential surfaces as a useful qualitative guide. Proc Natl Acad Sci U S A 93(20):10566-10571.
2. Pless SA, Millen KS, Hanek AP *et al* (2008) A Cation-pi Interaction in the Binding Site of the Glycine Receptor Is Mediated by a Phenylalanine Residue. J Neurosci 28(43):10937-10942.
3. Zhou Y, Xi Z, Chen W, Wang D (2008) Dinickel(II) Complexes of Bis(N-heterocyclic carbene) Ligands Containing Ni-2(mu-OH) Cores as Highly Efficient Catalysts for the Coupling of Aryl Chlorides. Organometallics 27(22):5911-5920.
4. Dougherty DA (2007) Cation-pi interactions involving aromatic amino acids. J Nutr 137(6):1504S-1508S.
5. Schottel BL, Chifotides HT, Dunbar KR (2008) Anion-pi interactions. Chem Soc Rev 37(1):68-83.
6. Burley SK, Petsko GA (1986) Amino-aromatic interactions in proteins. FEBS Lett 203(2):139-143.
7. Raines DE, Gioia F, Claycomb RJ, Stevens RJ (2004) The N-methyl-D-aspartate receptor inhibitory potencies of aromatic inhaled drugs of abuse: Evidence for modulation by cation-pi interactions. J Pharmacol Exp Ther 311(1):14-21.
8. Borissow CN, Black SJ, Paul M *et al* (2005) Adenophostin A and analogues modified at the adenine moiety: synthesis, conformational analysis and biological activity. Org Biomol Chem 3(2):245-252.
9. Zhu X-L, Xiong L, Li H *et al* (2014) Computational and Experimental Insight into the Molecular Mechanism of Carboxamide Inhibitors of Succinate-Ubquinone Oxidoreductase. Chemmedchem 9(7):1512-1521.
10. Crowley PB, Golovin A (2005) Cation-pi interactions in protein-protein interfaces. Proteins 59(2):231-239.
11. Scrutton NS, Raine AR (1996) Cation-pi bonding and amino-aromatic interactions in the biomolecular recognition of substituted ammonium ligands. Biochem J 319 (Pt 1):1-8.
12. Gallivan JP, Dougherty DA (1999) Cation-pi interactions in structural biology. Proc Natl Acad Sci U S A 96(17):9459-9464.
13. Liang Z, Li QX (2018) pi-Cation Interactions in Molecular Recognition: Perspectives on Pharmaceuticals and Pesticides. J Agric Food Chem 66(13):3315-3323.
14. Pellequer JL, Zhao B, Kao HI *et al* (2000) Stabilization of bound polycyclic aromatic hydrocarbons by a

pi-cation interaction. J Mol Biol 302(3):691-699.

15. Meyer EA, Castellano RK, Diederich F (2003) Interactions with aromatic rings in chemical and biological recognition. Angew Chem Int Ed 42(11):1210-1250.

16. Salonen LM, Ellermann M, Diederich F (2011) Aromatic Rings in Chemical and Biological Recognition: Energetics and Structures. Angew Chem Int Ed 50(21):4808-4842.

17. Hao G-F, Jiang W, Ye Y-N *et al* (2016) ACFIS: a web server for fragment-based drug discovery. Nucleic Acids Res 44(W1):W550-W556.

18. Wendler K, Thar J, Zahn S, Kirchner B (2010) Estimating the Hydrogen Bond Energy. J Phys Chem A 114(35):9529-9536.

19. Chourasia M, Sastry GM, Sastry GN (2011) Aromatic-Aromatic Interactions Database, A(2)ID: An analysis of aromatic pi-networks in proteins. Int J Biol Macromol 48(4):540-552.

20. Du Q-S, Long S-Y, Meng J-Z, Huang R-B (2012) Empirical Formulation and Parameterization of Cation-pi Interactions for Protein Modeling. J Comput Chem 33(2):153-162.

21. Salentin S, Schreiber S, Haupt VJ *et al* (2015) PLIP: fully automated protein–ligand interaction profiler. Nucleic Acids Res 43(W1):W443-W447.

22. Monera OD, Sereda TJ, Zhou NE *et al* (1995) Relationship of sidechain hydrophobicity and alpha-helical propensity on the stability of the single-stranded amphipathic alpha-helix. J Pept Sci 1(5):319-329.

23. Egner U, Hillig RC (2008) A structural biology view of target drugability. Expert Opin Drug Dis 3(4):391-401.

24. Settimo L, Bellman K, Knegtel RMA (2014) Comparison of the Accuracy of Experimental and Predicted pKa Values of Basic and Acidic Compounds. Pharm Res 31(4):1082-1095.

25. Schmidtke P, Le Guilloux V, Maupetit J, Tuffery P (2010) fpocket: online tools for protein ensemble pocket detection and tracking. Nucleic Acids Res 38:W582-W589.

26. Wang M-y, Wang F, Hao G-F, Yang G-F (2019) FungiPAD: A Free Web Tool for Compound Property Evaluation and Fungicide-Likeness Analysis. J Agric Food Chem 67(7):1823-1830.

27. Wu F-X, Wang F, Yang J-F *et al* (2018) AIMMS suite: a web server dedicated for prediction of drug resistance on protein mutation. Brief Bioinformatics 21(1):318-328.

28. Case DA, Cheatham TE, Darden T *et al* (2005) The Amber biomolecular simulation programs. J Comput

Chem 26(16):1668-1688.

29. Maier JA, Martinez C, Kasavajhala K *et al* (2015) ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. J Chem Theory Comput 11(8):3696-3713.

30. Wang JM, Wolf RM, Caldwell JW *et al* (2004) Development and testing of a general amber force field. J Comput Chem 25(9):1157-1174.

31. Wang F, Wu F-X, Li C-Z *et al* (2019) ACID: a free tool for drug repurposing using consensus inverse docking strategy. J Cheminformatics 11(1):73.

32. Genheden S, Ryde U (2015) The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. Expert Opin Drug Dis 10(5):449-461.

33. Krone MW, Albanese KI, Leighton GO *et al* (2020) Thermodynamic consequences of Tyr to Trp mutations in the cation–π-mediated binding of trimethyllysine by the HP1 chromodomain. Chem Sci 11(13):3495-3500.

34. Li H-L, Ma Y, Ma Y *et al* (2017) The design of novel inhibitors for treating cancer by targeting CDC25B through disruption of CDC25B-CDK2/Cyclin A interaction using computational approaches. Oncotarget 8(20):33225-33240.

35. Gallivan JP, Dougherty DA (1999) Cation-pi interactions in structural biology. Proc Natl Acad Sci U S A 96(17):9459-9464.

36. Dougherty DA (2013) The cation-π interaction. Acc Chem Res 46(4):885-893.

37. Guo Z, Li B, Cheng L-T *et al* (2015) Identification of Protein–Ligand Binding Sites by the Level-Set Variational Implicit-Solvent Approach. J Chem Theory Comput 11(2):753-765.

38. Wang F, Yang J-F, Wang M-Y *et al* (2020) Graph attention convolutional neural network model for chemical poisoning of honey bees' prediction. Sci Bull.

39. Bandyopadhyay D, Bhatnagar A, Jain S, Pratyaksh P (2020) Selective Stabilization of Aspartic Acid Protonation State within a Given Protein Conformation Occurs via Specific "Molecular Association". J Phys Chem B 124(26):5350-5361.

40. Choi H, Kang H, Park H (2013) New solvation free energy function comprising intermolecular solvation and intramolecular self-solvation terms. J Cheminformatics 5:8.

41. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (2012) Experimental and computational approaches to

estimate solubility and permeability in drug discovery and development settings. Adv Drug Del Rev 64:4-17.

42. Mahadevi AS, Sastry GN (2013) Cation-pi Interaction: Its Role and Relevance in Chemistry, Biology, and Material Science. Chem Rev 113(3):2100-2138.

43. Berry BW, Elvekrog MM, Tommos C (2007) Environmental modulation of protein cation-pi interactions. J Am Chem Soc 129(17):5308-5309.

44. Larrucea J, Rezabal E, Marino T *et al* (2010) Ab Initio Study of Microsolvated Al3+-Aromatic Amino Acid Complexes. J Phys Chem B 114(27):9017-9022.

45. Biot C, Buisine E, Rooman M (2003) Free-energy calculations of protein-ligand cation-pi and amino-pi interactions: From vacuum to proteinlike environments. J Am Chem Soc 125(46):13988-13994.

46. Xu YC, Shen JH, Zhu WL *et al* (2005) Influence of the water molecule on cation-pi interaction: Ab initio second order Moller-Plesset perturbation theory (MP2) calculations. J Phys Chem B 109(12):5945-5949.

47. Remko M, Soralova S (2012) Effect of water coordination on competition between pi and non-pi cation binding sites in aromatic amino acids: L-phenylalanine, L-tyrosine, and L-tryptophan Li+, Na+, and K+ complexes. J Biol Inorg Chem 17(4):621-630.

48. Aronov AM, Baker C, Bemis GW *et al* (2007) Flipped Out: Structure-Guided Design of Selective Pyrazolylpyrrole ERK Inhibitors. J Med Chem 50(6):1280-1287.

49. Sandanayaka V, Mamat B, Mishra RK *et al* (2010) Discovery of 4-[(2S)-2-{[4-(4-Chlorophenoxy)phenoxy]methyl}-1-pyrrolidinyl]butanoic Acid (DG-051) as a Novel Leukotriene A4 Hydrolase Inhibitor of Leukotriene B4 Biosynthesis. J Med Chem 53(2):573-585.

Figure legends

Fig. 1 Definition of cation-π and π-cation interactions. For the cation-π pair, the π system is typically provided by the aromatic side-chains of Trp, Tyr, and Phe, while the protonated ligand exists as the cation (PDB code: 1ax9). In the π-cation system, protonated Arg and Lys always act as cations with the arene ligands acting as a π partner (PDB code: 1bkm).

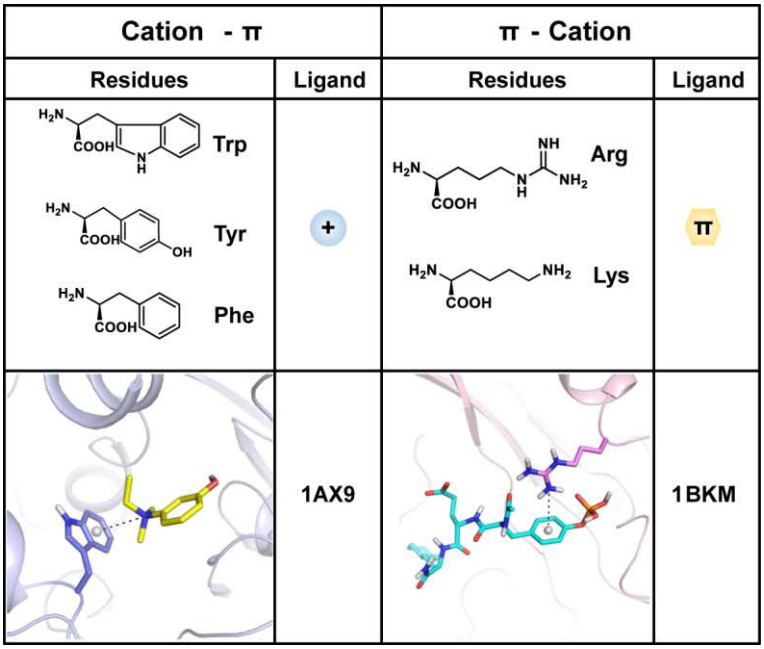
Fig. 2 (a) The workflow of screening cation-π and π-cation systems and (b) the number of residue pairs found in cation-π and π-cation systems. First, 141706 X-ray structures were downloaded from the PDB database. Then, approximately 60000 pdb files with small molecules were filtered out. Thirdly, 8263 complex crystal structures containing cation-π or π-cation interactions were screened by PLIP package. Finally, 1334 complexes with 1926 receptor-ligand cation-π interactions and 2174 complexes involved in 2643 receptor-ligand π-cation interactions were collected and classified according to the ligand groups.

Fig. 3 The histogram (a) and violin plot (b) of protein pocket hydrophobicity score in cation-π and π-cation systems.

Fig. 4 The histogram (a) and violin plot (b) of ligand pK_a in cation-π and π-cation systems.

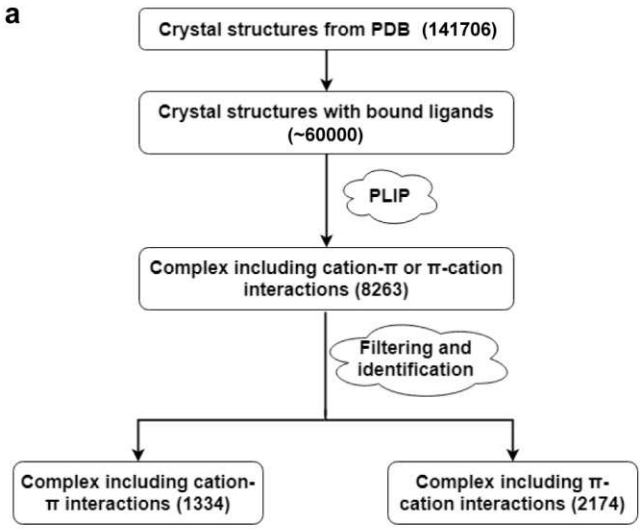
Fig. 5 The histogram (a) and violin plot (b) of SOL_Value in cation-π and π-cation systems.

Fig. 6 (a) Depicts of cation-π and π-cation interactions and (b) comparison of cation-π and π-cation interactions for IC_{50}/K_i, hydrophobic score (Hyd score), and SOL_Value.



416

417 Fig 1.

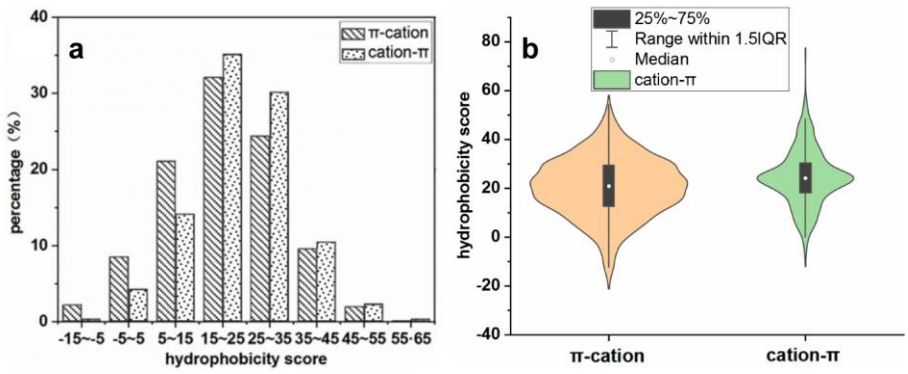


b The number of residue pairs in cation-π and π-cation systems

Cation-π (1334)			π-cation (2174)	
Trp	Tyr	Phe	Arg	Lys
1275	327	324	1782	861

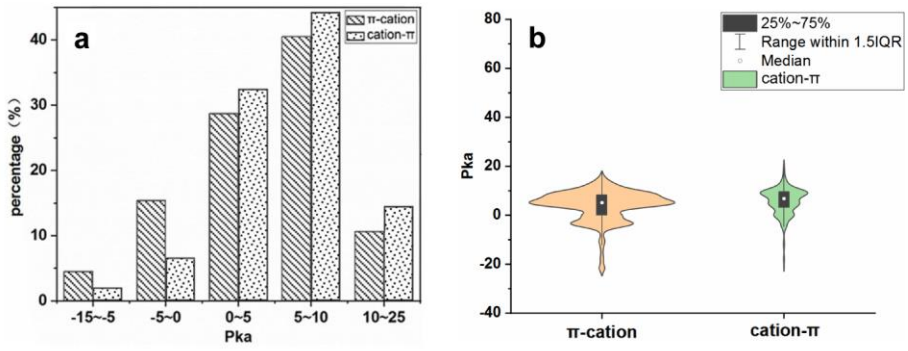
418

419 Fig 2.



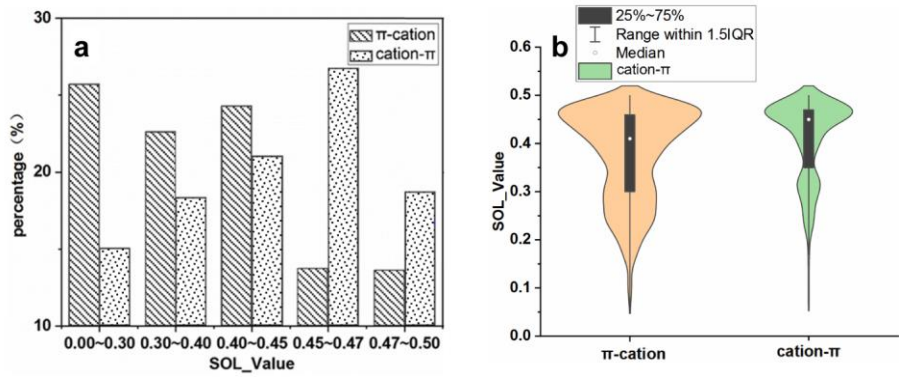
420

421 Fig 3.



422

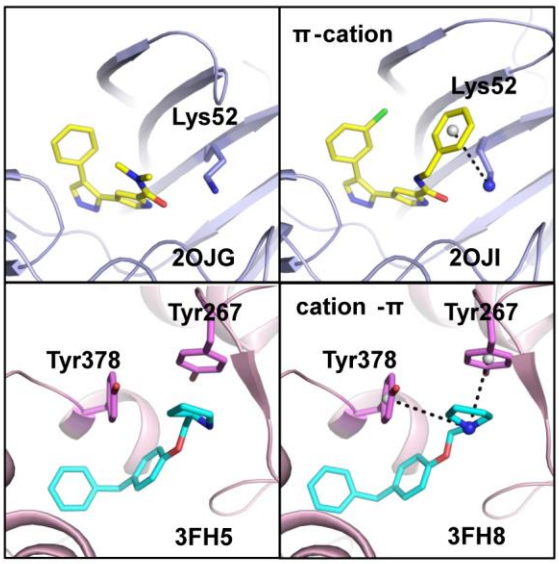
423 Fig 4.



424

425 Fig 5.

a



b The comparison of cation -π and π-cation interaction

PDB	IC_{50}/K_i(nM)	Hyd score	SOL_Value	π	cation
2OJG	2300	16.39	0.15		
2OJI	86	16.39	0.18	33A	Lys52
3FH5	87	42.41	0.39		
3FH8	26	42.41	0.41	Tyr267/378	27P

Fig 6.

Figures

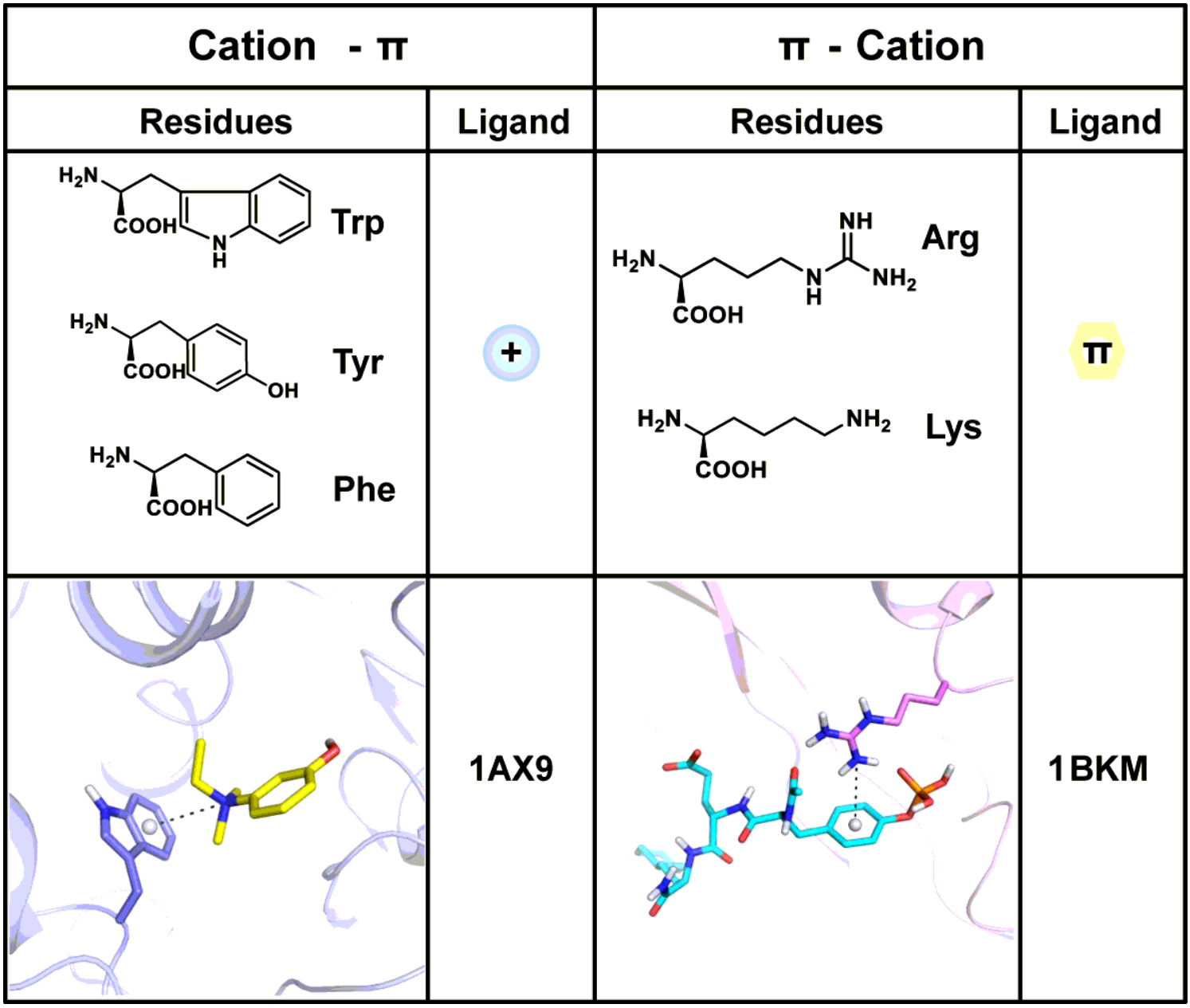
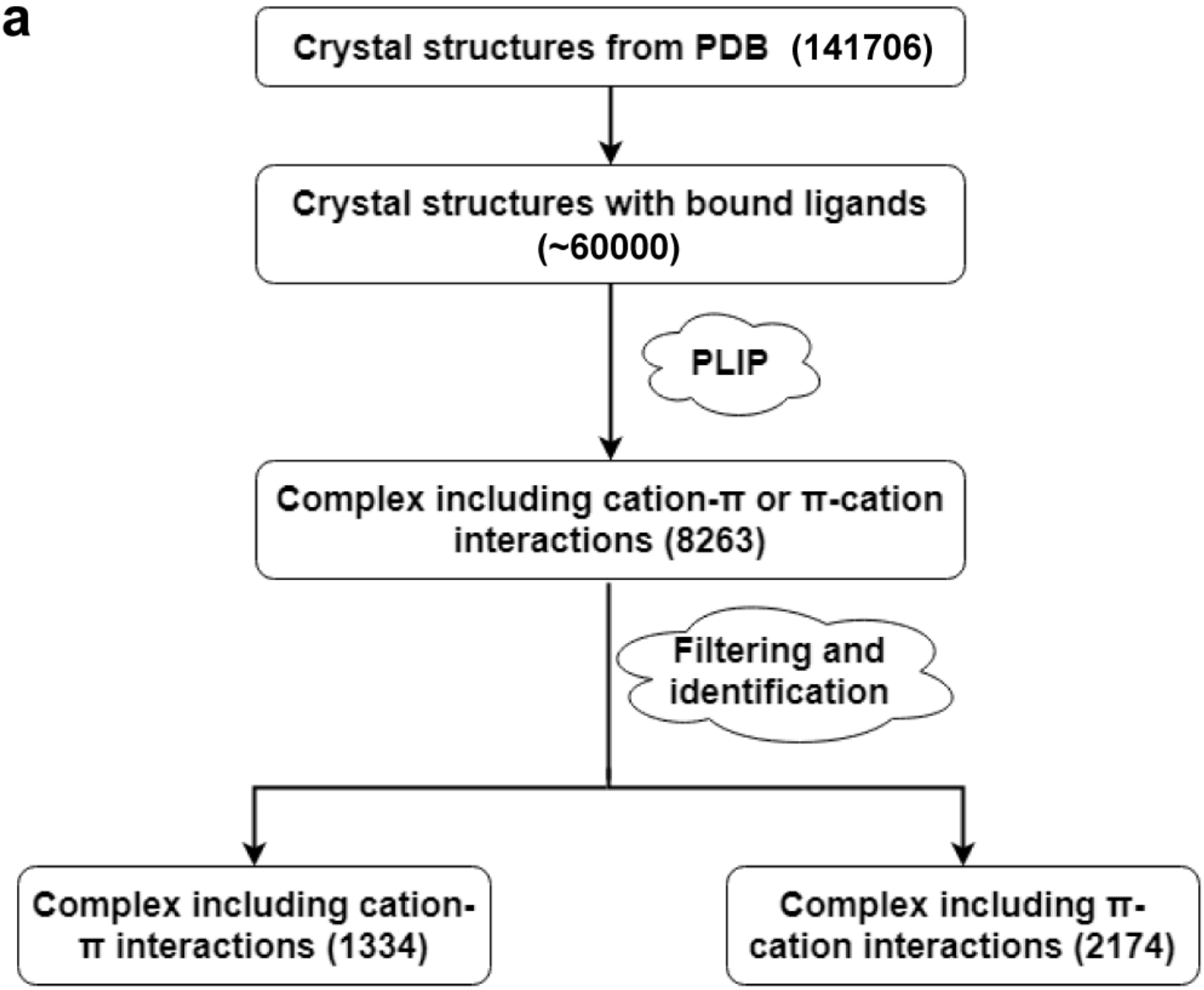


Figure 1

Definition of cation-π and π-cation interactions. For the cation-π pair, the π system is typically provided by the aromatic side-chains of Trp, Tyr, and Phe, while the protonated ligand exists as the cation (PDB code: 1ax9). In the π-cation system, protonated Arg and Lys always act as cations with the arene ligands acting as a π partner (PDB code: 1bkm).



b **The number of residue pairs in cation-π and π-cation systems**

Cation-π (1334)			π-cation (2174)	
Trp	Tyr	Phe	Arg	Lys
1275	327	324	1782	861

Figure 2

(a) The workflow of screening cation-π and π-cation systems and (b) the number of residue pairs found in cation-π and π-cation systems. First, 141706 X-ray structures were downloaded from the PDB database. Then, approximately 60000 pdb files with small molecules were filtered out. Thirdly, 8263 complex crystal structures containing cation-π or π-cation interactions were screened by PLIP package.

Finally, 1334 complexes with 1926 receptor-ligand cation-π interactions and 2174 complexes involved in 2643 receptor-ligand π-cation interactions were collected and classified according to the ligand groups.

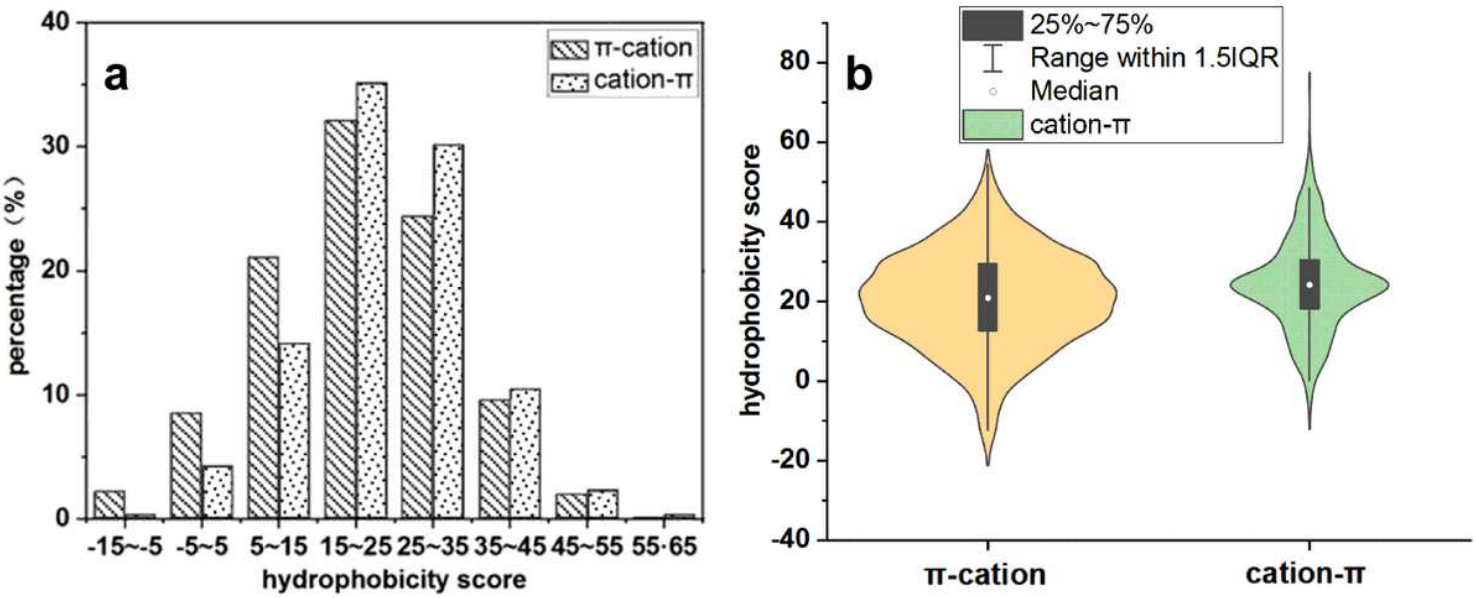


Figure 3

The histogram (a) and violin plot (b) of protein pocket hydrophobicity score in cation-π and π-cation systems.

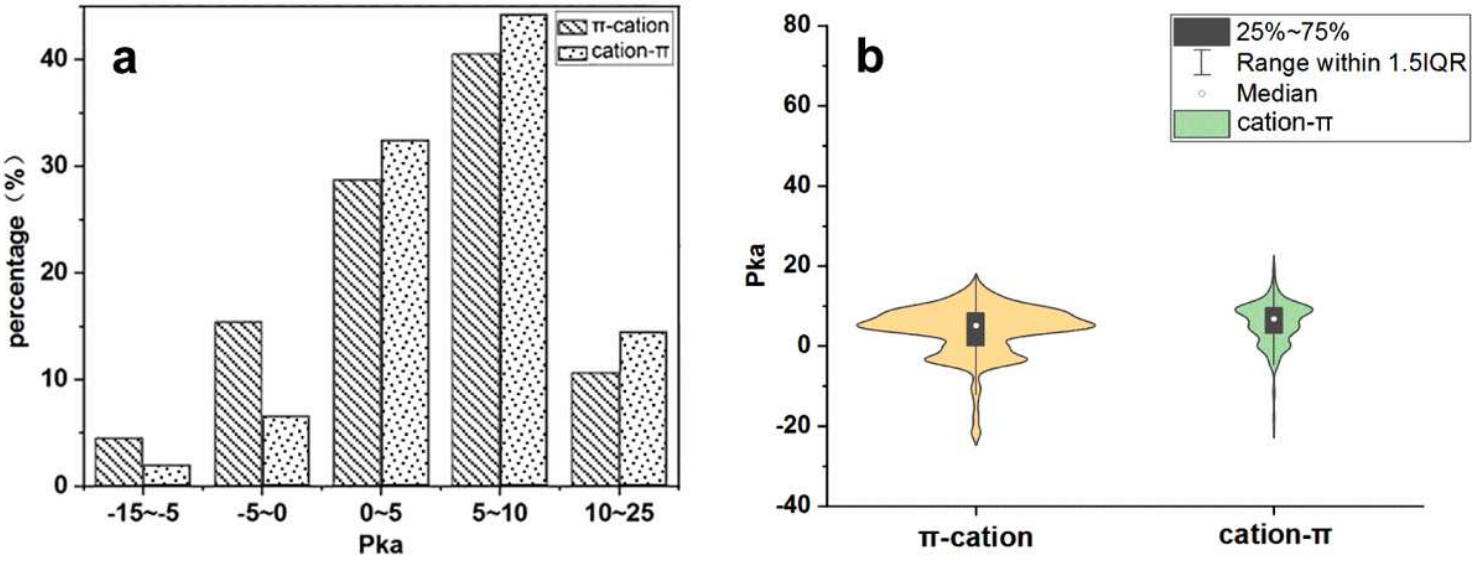


Figure 4

The histogram (a) and violin plot (b) of ligand pKa in cation-π and π-cation systems.

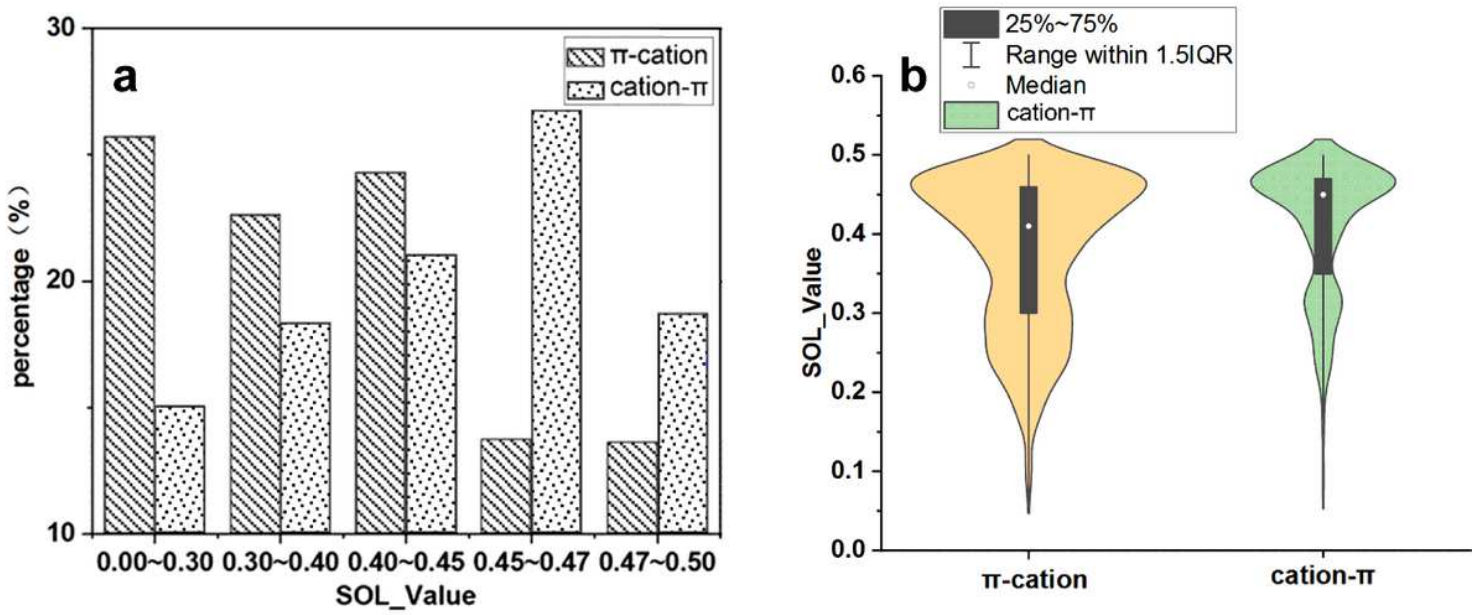
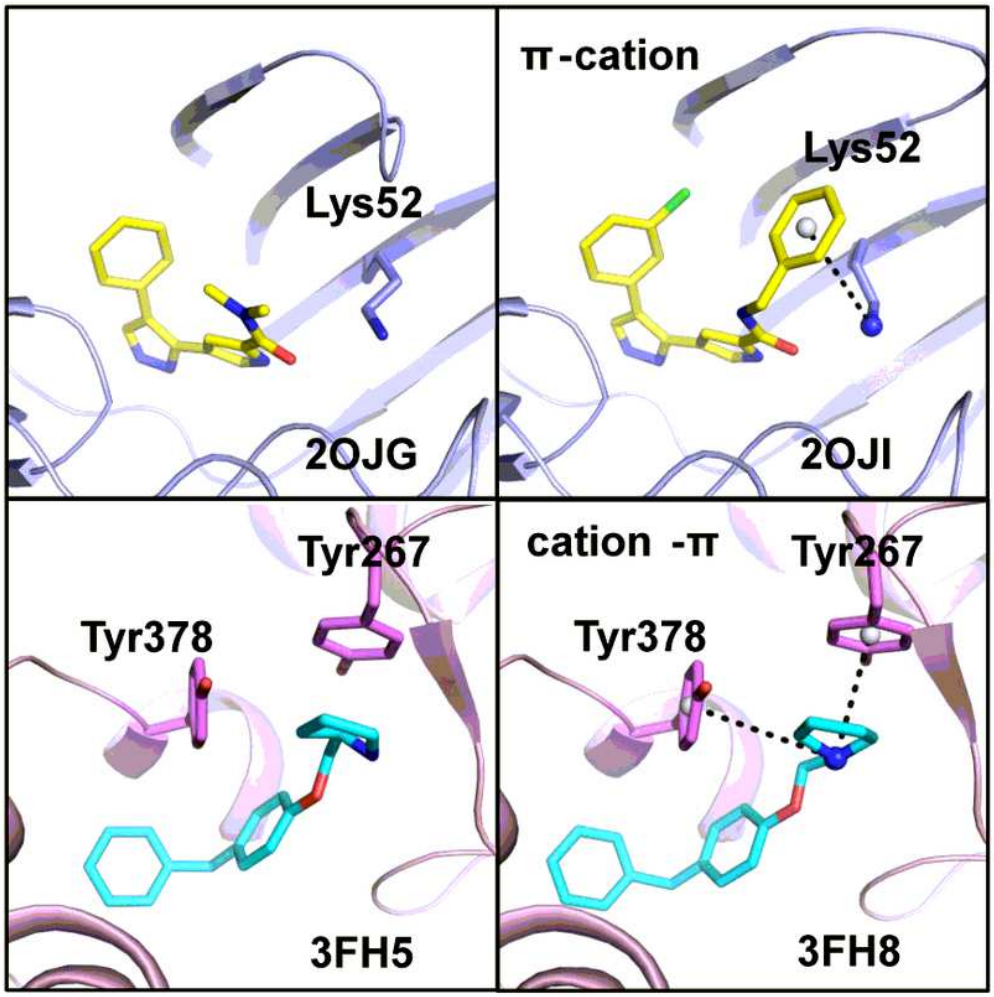


Figure 5

The histogram (a) and violin plot (b) of SOL_Value in cation-π and π-cation systems.

a



b The comparison of cation -π and π-cation interaction

PDB	IC_{50}/K_i(nM)	Hyd score	SOL_Value	π	cation
2OJG	2300	16.39	0.15		
2OJI	86	16.39	0.18	33A	Lys52
3FH5	87	42.41	0.39		
3FH8	26	42.41	0.41	Tyr267/378	27P

Figure 6

(a) Depicts of cation-π and π-cation interactions and (b) comparison of cation-π and π-cation interactions for IC50/Ki, hydrophobic score (Hyd score), and SOL_Value.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplementarymaterial20200818.doc