

A Systematic Analysis of Atomic Protein-Ligand Interactions in the PDB

Renato Ferreira de Freitas¹, Matthieu Schapira^{1,2,*}

¹ Structural Genomics Consortium, University of Toronto, Toronto, ON M5G 1L7, Canada

² Department of Pharmacology and Toxicology, University of Toronto, Toronto, ON M5S 1A8, Canada

*Correspondence should be addressed to matthieu.schapira@utoronto.ca

ABSTRACT

As the protein databank (PDB) recently passed the cap of 123,456 structures, it stands more than ever as an important resource not only to analyze structural features of specific biological systems, but also to study the prevalence of structural patterns observed in a large body of unrelated structures, that may reflect rules governing protein folding or molecular recognition. Here, we compiled a list of 11,016 unique structures of small-molecule ligands bound to proteins – 6,444 of which have experimental binding affinity - representing 750,873 protein-ligand atomic interactions, and analyzed the frequency, geometry and impact of each interaction type. We find that hydrophobic interactions are generally enriched in high-efficiency ligands, but polar interactions are over-represented in fragment inhibitors. While most observations extracted from the PDB will be familiar

to seasoned medicinal chemists, less expected findings, such as the high number of C–H
••O hydrogen bonds or the relatively frequent amide- π stacking between the backbone
amide of proteins and aromatic rings of ligands, uncover underused ligand design
strategies.

Introduction

Significant progress in high-throughput X-ray crystallography^{1,2} combined with advances in structural genomics³⁻⁵ have led to an explosion in the number of structures publicly available in the Protein Data Bank (PDB).⁶ At the time this manuscript was written, more than 123,456 structures had been deposited in the PDB⁶, including 76,056 protein-small molecule complexes, of which 13,000 have a reported binding potency.^{7,8} This large body of data contains important information on the nature, geometry, and frequency of atomic interactions that drive potent binding between small molecule ligands and their receptors. Systematic analysis of this data will lead to a better appreciation of intermolecular interactions between proteins and their ligands and can inform structure-based design and optimization of drugs.⁹

Several approaches have been developed for large-scale analysis of protein–small molecule interactions, such as SuperStar, or that implemented to build the Relibase database.^{10,11} PDBeMotif¹² and the recently published PELIKAN¹³ are two examples of free tools that can search for patterns in large collections of protein–ligand interfaces. Structural interaction fingerprints (SIF)⁹ are another method of representing and analyzing 3D protein-ligand interactions where the presence or absence of interactions between distinct residues and ligand atoms are represented as bit strings that can be compared rapidly.¹⁴ In addition, there has been an increase in the number of free tools to fully automate the detection and visualization of relevant non-covalent protein–ligand contacts in 3D structures.¹⁵⁻¹⁷

A statistical analysis of the nature, geometry and frequency of atomic interactions between small molecule ligands and their receptors in the PDB could inform the rational optimization of chemical series, help in the interpretation of difficult SAR, and serve as a knowledge-base for the improvement of scoring functions used in virtual screening. To the best of our knowledge, such public resource is currently missing.

Here, we analyze the frequency of common atomic interactions between protein and small molecules observed in the PDB. We find that some interactions occur more frequently in fragments than drug-like compounds, or in high-efficiency ligands than low-efficiency ligands. We next review in detail each of the most frequent interactions and use matched molecular pairs to illustrate the impact of these atomic interactions on binding affinity.

Most Frequent Protein-Ligand Atomic Interactions

We extracted from the PDB all X-ray structures of small-molecules in complex with proteins, with a resolution ≤ 2.5 Å, resulting in a collection of 11,016 complexes (only one structure was kept for each ligand). This collection contained 750,873 ligand-protein atom pairs, where a pair of atoms is defined as two atoms separated by 4Å or less. The top-100 most frequent ligand-protein atom pairs (Supplementary Table 1) can be clustered into seven interaction types (Figure 1). Among the most frequently observed are interactions that are well known and widely used in ligand design such as hydrophobic contacts, hydrogen bonds and π -stacking.^{18,19} These are followed by weak hydrogen bonds, salt bridges, amide stacking, and cation- π interactions.

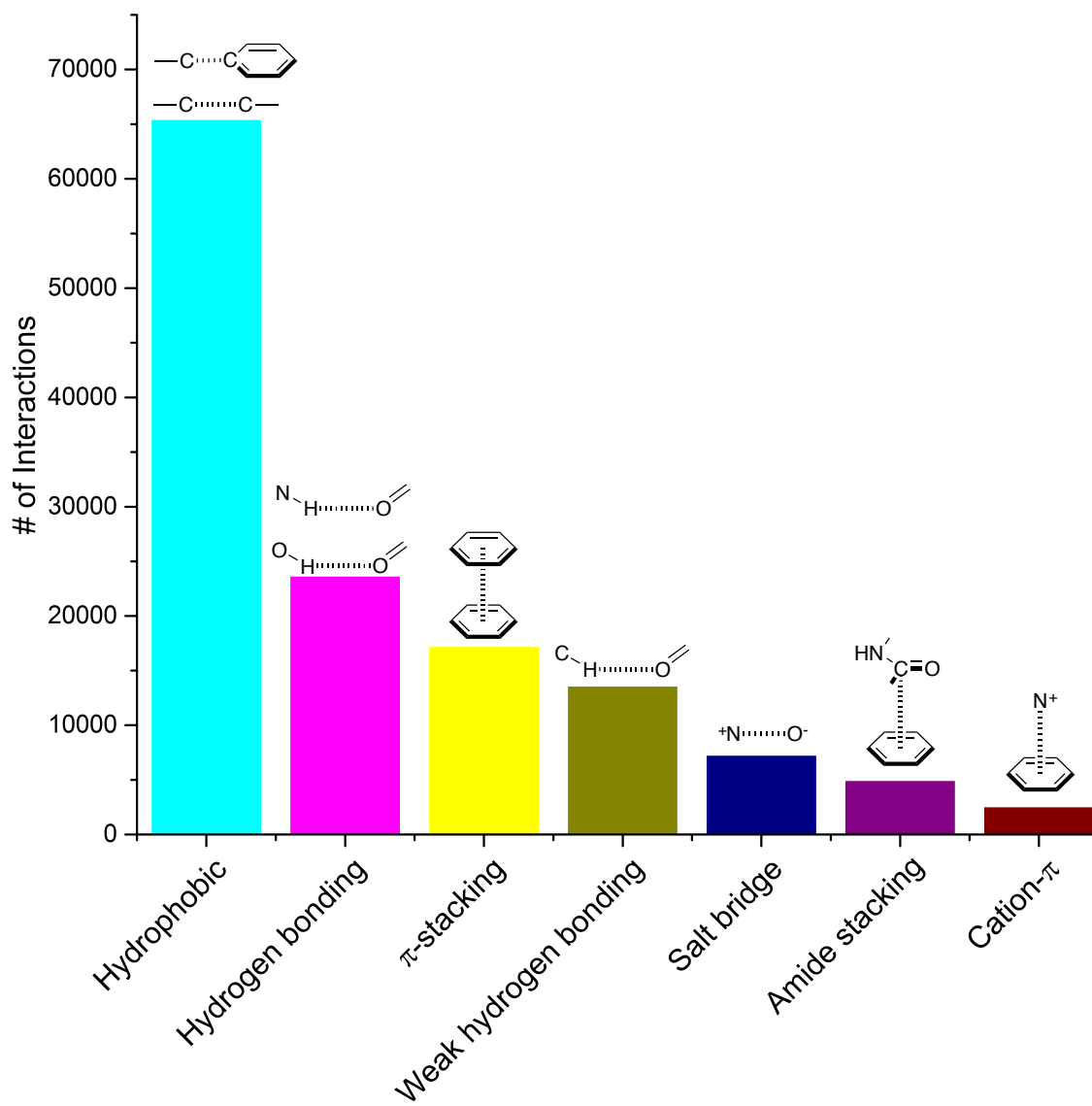


Figure 1. Frequency distribution of the most common non-covalent interactions observed in protein-ligands extracted from the PDB.

We next asked whether some interaction types were more frequently observed in high-efficiency ligands. Experimental binding affinity for 6444 protein-ligands in the PDB were retrieved from the PDBbind database^{7,8}, and a fit quality (FQ) score - a size-adjusted calculation of ligand efficiency - was used to evaluate how optimally a ligand binds relative to other ligands of any size.²⁰ The frequency of each interaction type was

calculated in the 1500 protein-ligand complexes with the best FQ score (FQ > 0.81) and the 1500 complexes with the worst FQ scores (FQ < 0.54) (Figure 2a).

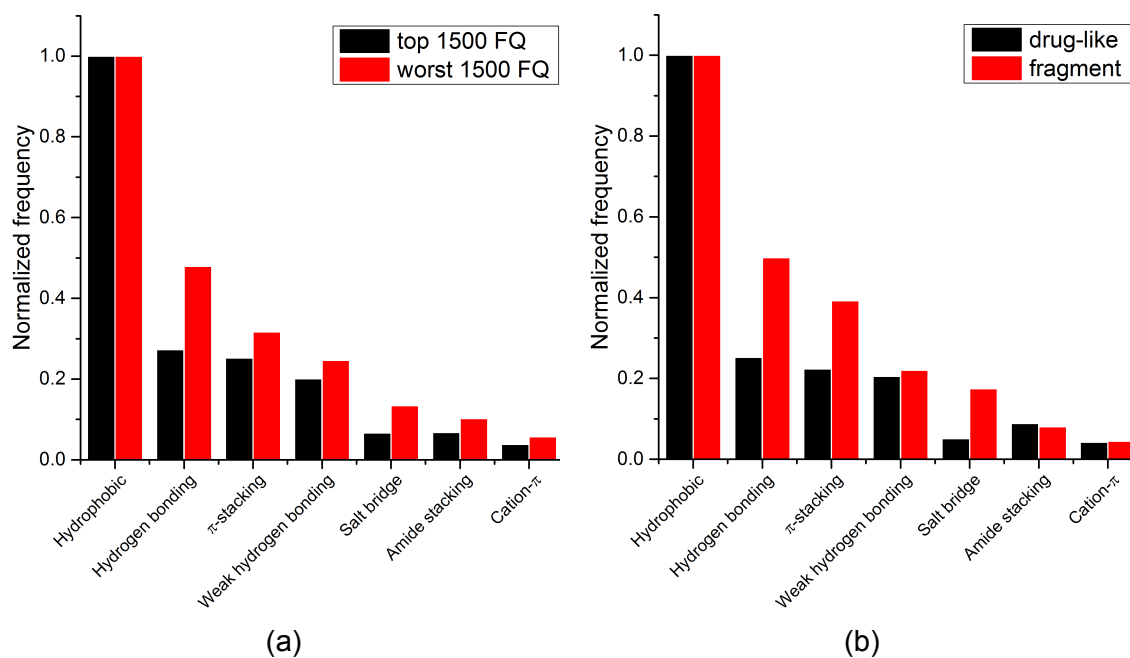


Figure 2. Relative frequency distribution of the most common non-covalent interactions observed in: (a) ligands with high vs. low ligand efficiency; (b) fragments vs. drug-like compounds.

We find that hydrophobic interactions are more frequent in high-efficiency ligands. In particular, the relative contribution of hydrogen bonds is reduced from 48% to 27% of that of hydrophobic contacts in efficient binders, and the contribution of salt bridges is more than halved, from 13% to 7%. This observation probably reflects the fact that most ligands in the PDB are the product of lead optimization strategies that aim at increasing the number of favorable hydrophobic interactions, which is less challenging than optimizing directionality-constrained hydrogen-bonds (discussed in more details in a later section).²¹ We also find that efficient ligands are more hydrophobic, as the median

number of heavy atoms and logD (ChemAxon) for compounds with high FQ are 27 and 1.7 respectively, and 21 and 0.2 respectively for compounds with low FQ. Both groups showed similar profiles for other properties like polar surface area (median PSA: 95.3 vs. 89.8 Å²), hydrogen bond acceptors (median HBA: 5 for both), and hydrogen bond donors (median HBD: 2 for both). Taken together these results show that small-molecule ligands that bind their target with high efficiency are more hydrophobic, and that hydrophobic interactions are a driving factor for increased ligand efficiency.

Since fragments are typically binding their targets with higher efficiency than larger ligands, we asked whether hydrophobic interactions were also more frequent in protein-fragment complexes. The frequency of each interaction type was calculated for two random groups of 1500 hundred protein-ligand complexes, one with fragment molecules ($HA \leq 20$), the other with drug-like molecules ($30 \leq HA \leq 50$) (Figure 2b). Unlike high-efficiency ligands, we find that protein-fragment complexes are enriched in polar interactions: the frequency of hydrogen bond is doubled, from 25% to 50%, compared to that of hydrophobic contacts, and the frequency of electrostatic interactions is multiplied by three, from 5% to 17%. To compensate for their low number, interactions made by fragments need to be highly efficient.²² We note that electrostatic interactions define maximum efficiency of ligand binding.²³ Similarly, though to a lesser extent, optimal hydrogen-bonds are energetically more favorable than hydrophobic contacts.²⁴ The observed enrichment in polar interactions reflects the fact that fragments are freer than chemical moieties of larger compounds to adopt binding poses that will optimally satisfy the geometric constraints of high-efficiency interactions, such as electrostatic or hydrogen-bonds.²⁵

Together, these results show that fragments are using polar interactions to gain maximum binding efficiency from a limited number of interactions, but as small-molecule ligands are optimized, geometric constraints associated with polar bonds are more challenging to satisfy, and the contribution of hydrophobic interactions increases. To gain further insight, we next analyzed in detail the composition, geometry, frequency, protein side-chain preference, and impact towards binding affinity of each protein-ligand interaction type in the PDB.

Specific Intermolecular Interactions

Hydrophobic interactions

From our analysis, hydrophobic contacts are by far the most common interactions in drug-receptor complexes, totalizing 66772 contacts between a carbon and a carbon, halogen or sulfur atom (the distance cut-off of 4.0 Å allows the implicit inclusion of hydrogen atoms) (Figure 1). Hydrophobic interactions were separated into five groups (Table S2). The most populated group is the one formed by an aliphatic carbon in the receptor and an aromatic carbon in the ligand, which alone accounts for more than 42000 interactions (Table S2). This is an indication that aromatic rings are prevalent in small molecule inhibitors. In fact, 76% of the marketed drugs contain one or more aromatic ring, with benzene being by far the most frequently encountered ring system.^{26,27} Not surprisingly, leucine, followed by valine, isoleucine and alanine side-chains are the most frequently engaged in hydrophobic interactions (Figure S2).

Contacts involving an aromatic or aliphatic carbon in the receptor and an aliphatic carbon in the ligand were observed in 8899 and 8974 instances, respectively (Table S2). We observed that aliphatic carbons were distributed mostly above or below the plane of the aromatic ring, rather than at the edge (Figure S3). Although we didn't investigate further the directionality of these interactions we believed that some of them could be further classified as C-H... π interactions.²⁸ These interactions occur ubiquitously in almost all proteins, making important contributions to biomolecular structures and ligand binding.^{29,30} Interactions involving an aliphatic or aromatic carbon in the protein and a chlorine or fluorine in the ligand were observed 5147 times (Table S2). While these atomic contacts are hydrophobic, a C-X... π component can also contribute to this type of interaction.^{31,32} Interactions involving a sulfur atom from the side chain of methionine and an aromatic carbon from the ligand were observed in 1309 complexes. Although methionine is classified as a hydrophobic residue, a recent study shows that the Met S...C(aromatic) interaction yields an additional stabilization energy of 1–1.5 kcal/mol compared with a purely hydrophobic interaction.³³

Hydrophobic interactions are the main driving force in drug-receptor interactions. The benefit of burying a solvent-exposed methyl group on a ligand into a hydrophobic pocket of a protein is about 0.7 kcal/mol or a 3.2-fold increase in binding constant per methyl group.³⁴ However the effect of replacing a hydrogen atom with a methyl group is highly context dependent, and potency losses are as common as gains. Ten-fold and 100-fold gains in potency are observed in 8% and 1% of cases respectively.^{35,36} For instance, addition of a single methyl group improves by 50 fold the potency of a tankyrase-2 (TNKS2) inhibitor.³⁷ The added methyl group occupies a small hydrophobic cavity and

potentially releases unfavorably bound water molecules (Figure 3b). In rare cases, the increase in potency due to the introduction of a “magic methyl” exceeds two orders of magnitude.³⁸ This is generally due to the combined entropic effect of lowering the conformational penalty paid by the ligand upon binding, and the desolvation effect of burying the methyl group in a hydrophobic pocket.³⁵

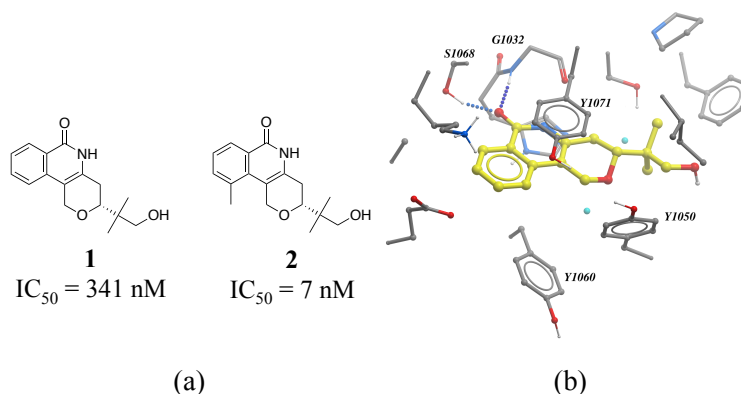


Figure 3. Magic methyl effect: (a) Chemical structure of two TNKS2 inhibitors; (b) crystal structure of **1** (carbon atoms in yellow) bound to TNKS2 (PDB: 5C5P). Oxygen atoms are in red (cyan for water molecules).

Hydrogen bonds

We find that hydrogen bonds were the second most frequent type of interactions observed in our collection of protein-ligand complexes, with a total number of 24058 instances (Figure 1). N-H...O interactions were more frequent (15105 interactions) than O-H...O (8251 interactions) and N-H...N (333 interactions) (Table S2). Among the N-H...O interactions, the number of neutral and charged hydrogen bonds were almost equal (7554 vs. 7551, respectively). Proteins were more often hydrogen-bond donors than

acceptors (9217 vs. 5888, respectively). Surprisingly, glycine was the most frequent hydrogen-bond acceptor, and the second most frequent donor (Figure S4). Arginines were engaged in more hydrogen-bonds than lysines, probably reflecting the presence of 3 nitrogen atoms in the guanidinium group of arginine side-chains (Figure S4). Among O-H...O interactions, charged hydrogen bonds (typically between an alcohol and a carboxylic acid) were 3 times more frequent than neutral ones, and ligands more often behaved as donors than acceptors (Table S2). The most common acceptors were aspartic acids in charged hydrogen bonds, and asparagine, glycine and glutamine in neutral interactions (Figure S5). Serine was the most usual donor (Figure S5).

We found that heavy atoms in N-H...O, N-H...N, and O-H...O hydrogen bonds were all separated by similar median distances of approximately 3.0 Å (Figure 4). This value is slightly higher (~ 0.1-0.2 Å) than previously reported for hydrogen bonds between amide C=O and OH/NH.³⁹ In addition, the median distances of neutral and charged hydrogen bonds were almost identical (0.1 Å difference, data not shown). The D-H...A angle usually peaked at 130-180°, and the preferred angle for N-H...O hydrogen bonds was around 180° (data not shown)

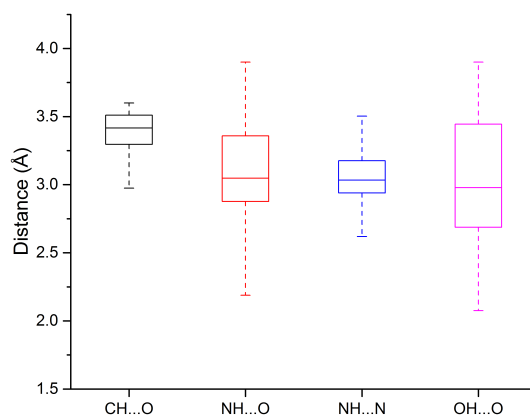


Figure 4. Box plots of hydrogen bond length distributions for the weak (C-H...O) and strong hydrogen bonds (N-H...O, N-H...N, O-H...O).

Hydrogen bonds are the prevailing directional intermolecular interactions in biological complexes⁴⁰, and the predominant contribution to the specificity of molecular recognition.⁴¹ Hydrogen-bond donors and acceptors not only play important roles in DNA base-pairing⁴², protein folding^{43,44}, and enzyme specificity⁴⁵, but also in modulating the physicochemical properties - in particular the lipophilicity⁴⁶ - of ligands, a key parameter in drug design. The free energy for hydrogen bonding can vary between -1.5 kcal/mol to -4.7 kcal/mol.³⁴ However, the contribution of a hydrogen bond to binding can be very modest (or penalizing) if the new interaction formed does not outweigh the desolvation penalty upon ligand binding.⁴⁷ Also, the contribution of a hydrogen bond is dependent on the local environment: a solvent-exposed hydrogen-bond contributes significantly less to net interaction energy than the same hydrogen-bond in a buried hydrophobic pocket.⁴⁸ Consequently, optimizing hydrophobic interactions is generally considered easier than hydrogen bonds.³⁴ In drug design, hydrogen bonds are exploited to gain specificity owing to their strict distance and geometric constraints.⁴⁹

Among numerous examples, a series of potent thrombin inhibitors shows a remarkable increase in binding affinity (> 500-fold) through simple addition of hydrogen-donating ammonium group (Figure 5).⁵⁰ In the crystal structure, the ammonium group forms a charge-assisted hydrogen bond with the carbonyl oxygen of Gly216 and surrounding waters.⁵¹

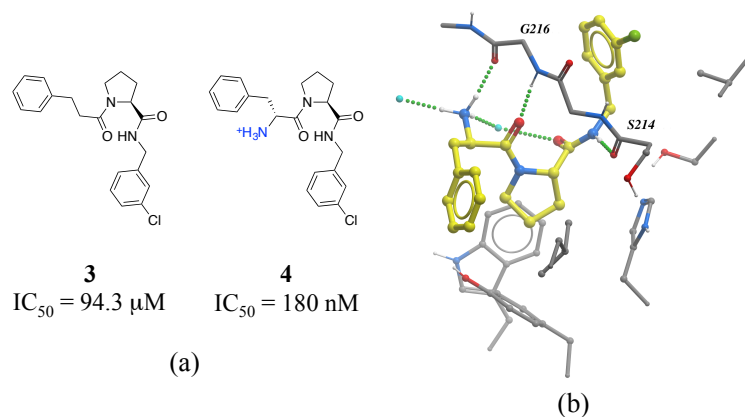


Figure 5. Effect of adding a hydrogen bond in a thrombin inhibitor: a) chemical structure of a pair of thrombin inhibitors; b) crystal structure of **4** (yellow carbons) in complex with thrombin (PDB: 2ZC9). Hydrogen bonds are displayed in dotted green lines and waters oxygens as cyan.

π -stacking interactions

The third most frequent protein-ligand contacts in the PDB were aromatic interactions (Figure. 1). Interactions involving aromatic rings are ubiquitous in chemical and biological systems and can be considered a special case of hydrophobic interactions.²⁸ We found that edge-to-face and face-to-face interactions were equiprobable (8704 and 8537 contacts respectively) (Table S2). This is in agreement with quantum mechanical calculations of the interaction energy of benzene dimers that predict

the edge-to-face and parallel displaced face-to-face as being isoenergetic, and more stable than the eclipsed face-to-face π -stacking.⁵² Almost 50% of all π -stacking interactions are observed between the aromatic ring of phenylalanine and an aromatic ring in the ligand, followed by tyrosine (36.8%), tryptophan (8.7%) and histidine (5.1%) (Figure S6).

Interactions involving aromatic rings are major contributors to protein–ligand recognition and concomitantly to drug design.^{28,53} An example of the strong gain in binding affinity that can be obtained by forming a π -stacking interaction is illustrated in a series of soluble epoxide hydrolase (sEH) inhibitors.⁵⁴ In the X-ray cocrystal structure of human sEH and **6** ($IC_{50} = 7$ nM), the phenyl ring is positioned to allow π -stacking interaction with H524 (Figure 6b), while an analog (**5**, $IC_{50} = 700$ nM) without the phenyl ring is 100-fold less potent. While π -stacking interaction can increase the binding affinity of the inhibitor for its target, it has been pointed out that reducing the number of aromatic rings of a molecule might improve its physicochemical properties, such as solubility.^{55,56}

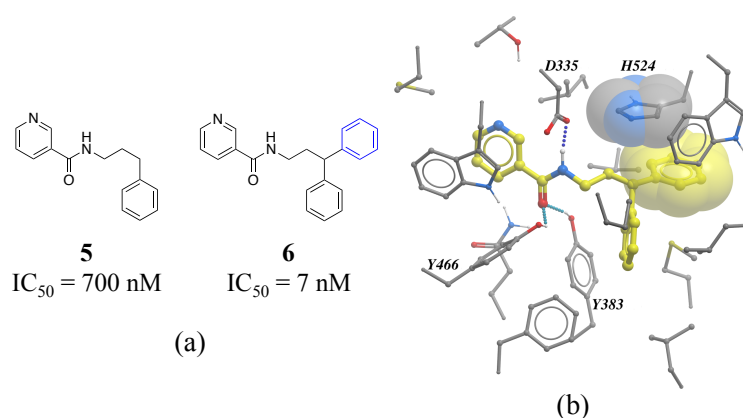


Figure 6. a) Chemical structure of two inhibitors of human sEH; b) X-ray cocrystal structure of human sEH and **6** (yellow carbons, PDB: 3I1Y), the phenyl ring is positioned to allow a π -stacking interaction with H524. Hydrogen bonds are displayed in dotted green lines.

Weak hydrogen bonds

The fourth most frequent interactions (13600 contacts) were C-H...O hydrogen bonds, the existence of which is well documented (Figure 1).^{57,58} When the interacting carbon was aromatic, protein oxygens were found to be acceptors much more often than ligand oxygen atoms (4927 vs. 708 interactions, Table S2). This simply reflects the fact that most ligands have aromatic rings, while most side-chains don't. Glycine, aspartic acid and glutamic acid were always the most frequent acceptors in C-H...O interactions, while leucine was the most frequent donor (Figure S7).

The median distance of the C-H...O hydrogen bonding was 3.4 Å, which is 0.4 Å longer than traditional hydrogen bonds (N-H...O, N-H...N, O-H...O), and distances separating the two heavy atoms were rarely lower than 3.2 Å (Figure 3). The angle distribution of C-H...O interactions peaked around 130° (data not shown), which is in agreement with previous work.⁵⁹

The existence of weak hydrogen bonds has been extensively analyzed and reviewed.⁶⁰⁻⁶³ Calculations indicate that the magnitude of the C_α-H...O=C interactions are about one-half the strength of an NH...O=C hydrogen bond.⁶⁴ In addition, an analysis of protein-ligand complexes revealed that C_α-H...O hydrogen should be better interpreted as secondary interactions, as they are frequently accompanied by bifurcated N-H...O hydrogen bonds.⁶⁵ However, it is increasingly recognized that C-H...O hydrogen bonds play an important role in molecular recognition processes⁶⁶, protein folding stabilization⁶⁷, in the interaction of nucleic acids with proteins⁶⁸, in enzyme catalysis⁶⁹, and in the stabilization of protein-ligand binding complexes.^{70,71} A matched pair of CDK2

inhibitors illustrates the contribution of C-H...O hydrogen bonds to protein-ligand complexes (Figure 7).⁷² The only difference between the two inhibitors is the substitution of a NH₂ with a methyl group on the thiazole ring of compound **8** (Figure 7a). Although the N-H...O hydrogen bond of **7** is stronger than the C-H...O hydrogen bond of **8**, the latter compound is more potent probably due to the penalty associated with desolvating the NH₂ of **7** upon binding.

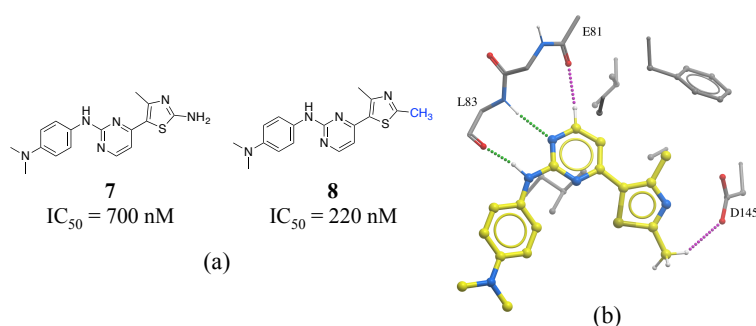


Figure 7. a) Chemical structure of two CDK2 inhibitors; b) X-ray cocrystal structure of human CDK2 and **8** (PDB: 1PXP). The N-H...O and C-H...O hydrogen bonds are displayed as green and magenta dotted lines, respectively.

Salt bridges

The contact between a positively charged nitrogen and a negatively charged oxygen (i.e. *salt bridge*) was the fifth most frequent interaction type in our analysis (7276 interactions) (Figure 1). The number of salt bridge interactions with a positive nitrogen coming from the protein and the negative oxygen coming from the ligand was two times higher than the opposite (4882 vs. 2394 interactions, Table S2). This probably reflects the higher number of ligands containing carboxylic acids (1849) than ammonium groups

(1103) in our dataset, as the frequency of arginine (5.6%) and lysine (5.0%) in proteins is similar to that observed for aspartic acid (5.4%) and glutamic acid (3.8%) (UniProtKB/TrEMBL UniProt release 2017_03).⁷³ Arginine was the cation in 83.6% of all interactions (Figure S8). This seems to be agreement with quantum mechanical calculations, which predict that arginine are more inclined than lysine side-chains to form salt bridges.⁷⁴ Finally, the distribution of negatively charged oxygens around the guanidinium group of arginine shows a higher density around the terminal (ω) nitrogens than at the secondary amine (ϵ) nitrogen (Figure S8).

Salt bridges contribute little to protein stability as the favorable binding energy obtained from forming a salt bridge is not sufficient to offset the energetic penalty of desolvating charged groups.^{75,76} However, the strength of salt bridge interactions is strongly dependent on the environment. In particular, buried salt-bridges can make crucial contributions to ligand binding.⁷⁷⁻⁷⁹ For example, the terminal N,N-dimethylamino tail of **10** forms a salt bridge with D831 in the kinase domain of epidermal growth factor receptor (EGFR) (Figure 8). When the nitrogen atom of the terminal N,N-dimethylamino group was replaced by a carbon (**9**) potency was reduced by more than 800-fold.⁸⁰

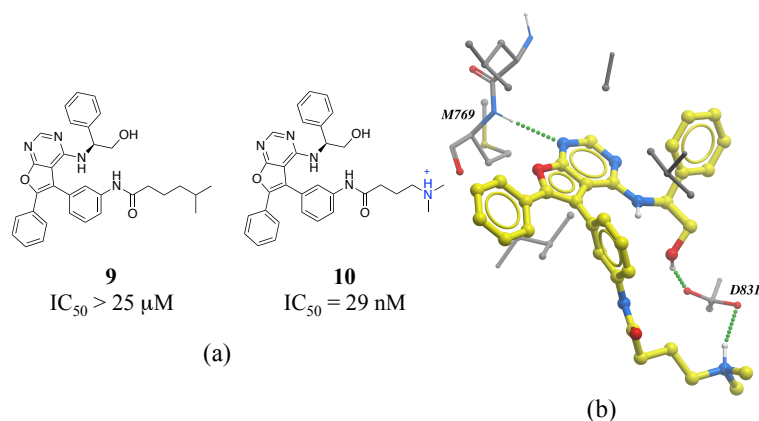


Figure 8. a) Chemical structure of two inhibitors of human EGFR; b) X-ray cocrystal structure of kinase domain of EGFR and **10** (PDB: 4JRV), the terminal N,N -dimethylamino tail of **10** forms a salt bridge with D831. Hydrogen bonds are displayed in dotted green lines.

Amide $\cdots\pi$ stacking

Interactions between an amide group and an aromatic ring were the sixth most frequently observed (Figure 1). In these interactions, which are related to canonical aromatic π -stacking, the π -surface of the amide bond stacks against the π -surface of the aromatic ring.^{81,82} As previously observed for π - π stacking interactions, we did not find significant preference for face-to-face over edge-to-face arrangement (2907 and 2060 interactions respectively) (Table S2). The most frequent amino acids participating in face-to-face amide $\cdots\pi$ stacking were glycine (19.4%) and tryptophan (17.9%), while glycine (20.1%) and leucine (13.0%) were the most often observed in edge-to-face geometry (Figure S9). The fact that 88.5% of all amide $\cdots\pi$ stacking interactions occurred between the backbone amide group of a protein (generally a glycine) and the aromatic ring of a ligand points at a strategy to exploit peptide bonds in binding sites that is probably underused in structure-based drug design.

Amide $\cdots\pi$ stacking interactions are common and significant in protein structures.⁸³ These interactions were also shown to sometimes play an important role in ligand binding.^{84–86} For example, the selective permeation of urea is facilitated by amide $\cdots\pi$ stacking interactions in the urea transporter.⁸⁷ In another example, the 11-fold difference in K_i between a matched pair of oxazole-containing factor Xa inhibitors was attributed to the influence of the dipole of the oxazole ring on the amide $\cdots\pi$ interaction (Figure 9).⁸⁸

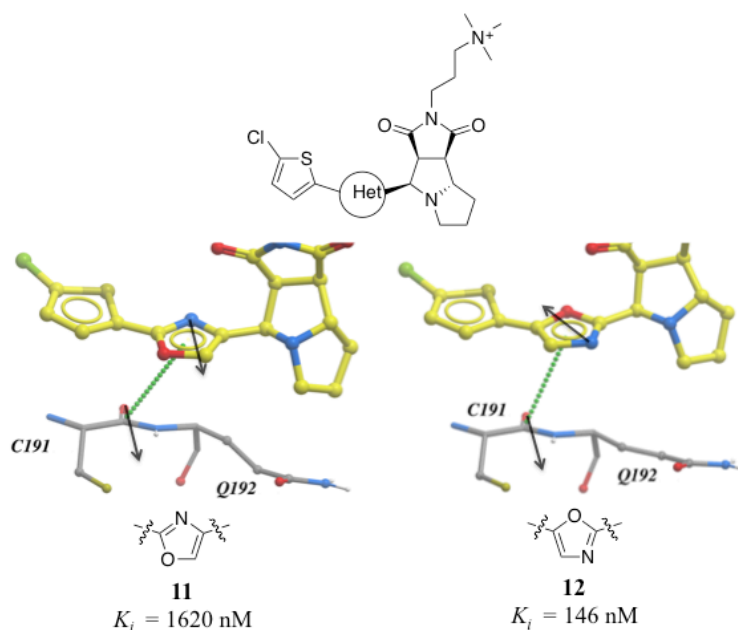


Figure 9. X-ray cocrystal structure of **11** (PDB: 2Y5H) and **12** (PDB: 2Y5G) bound at the active site of factor Xa. The amide $\cdots\pi$ stacking interaction is shown as dotted green lines. The dipoles of the oxazole ring and peptide amide (black arrows) are parallel in **11** and anti-parallel in **12**.

Cation- π

We found 2577 interactions between a positively charged nitrogen and an aromatic ring (Figure 1). These cation- π interactions are essentially electrostatic due to

the negatively charged electron cloud of π systems.⁸⁹ In more than 90% of these interactions, the nitrogen came from the receptor and the aromatic ring from the ligand, reflecting, as previously noted, that drug-like compounds have often aromatic rings while ammonium groups are more rare (Table S2). Arginines were 3 times more frequently engaged in cation- π interactions than lysine side-chains. (Figure S10). A similar trend was previously observed for peptidic interactions.⁹⁰ This preference has been attributed to the fact that the guanidinium group of arginines can donate several hydrogen bonds while simultaneously binding to an aromatic ring.⁸² When the positive nitrogen came from the ligand, tyrosine side-chains were the most common partner with 156 interactions, followed by phenylalanine and tryptophan (59 and 24 interactions respectively) (Figure S10). Potentiation of the cation- π binding ability of the tyrosine upon hydrogen bonding of its hydroxyl group was proposed to be at the origin of a similar bias in peptidic interactions.⁹⁰

Cation- π interactions are widespread in proteins and are important determinants of the structure, stability, and function of proteins.⁹¹ An example that is especially compelling is the Royal family of epigenetic reader proteins, that feature an aromatic cage composed of two to four aromatic residues that make cation- π and hydrophobic interactions with postranslationally methylated lysines or arginines side-chains.⁹²

Many drug-receptor interactions involve cation- π interactions. One of the earliest examples is the recognition of acetylcholine (ACh) by the nicotinic acetylcholine receptor (nAChR). Similarly, GABA⁹³, glycine⁹⁴, and 5-HT₃⁹⁵ receptors have all been shown to participate in cation- π interactions with neurotransmitters. In a series of insightful experiments, sequential methylation of an ammonium group in a series of

potent factor Xa inhibitors gradually increased the binding affinity by 3 orders of magnitude.⁹⁶ Comparing the affinity of a ter-butyl analog (compound **15**) with the trimethylated ammonium group (compound **17**), indicated that the cation- π interaction contributed to a 60 fold increase in potency (Figure 10).

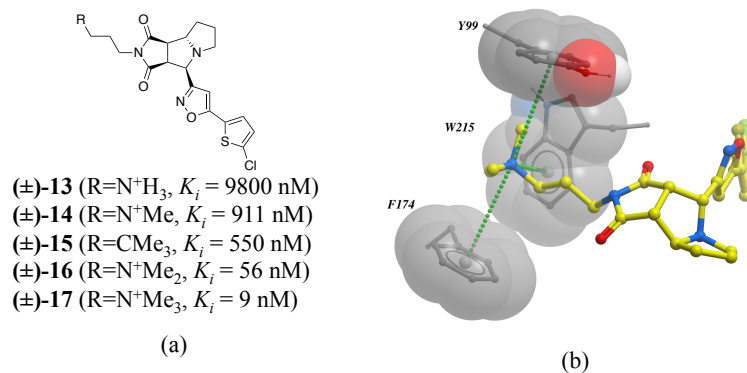


Figure 10. a) Chemical structure of a series inhibitors of human factor Xa; b) in the X-ray cocrystal structure of human factor Xa and **17** (PDB: 2JKH), the quaternary ammonium ion fill the aromatic box. The cation- π interaction is displayed as dotted green lines.

Conclusion

We presented here a statistical analysis of the nature, geometry and frequency of atomic interactions between small molecule ligands and their receptors available in the PDB. The enrichment of polar interactions in bound fragments, but hydrophobic contacts in optimized compounds reflects the challenge of overcoming desolvation penalty during lead optimization. This unbiased census recapitulates well-known rules driving ligand design, but also uncovers some interaction types that are often overlooked in medicinal chemistry. This analysis will help in the interpretation of difficult SAR, and may serve as a knowledgebase for the improvement of scoring functions used in virtual screening.

Acknowledgments

We thank Vijayaratnam Santhakumar for his helpful comments on this manuscript. The SGC is a registered charity (number 1097737) that receives funds from AbbVie, Bayer Pharma AG, Boehringer Ingelheim, Canada Foundation for Innovation, Eshelman Institute for Innovation, Genome Canada through Ontario Genomics Institute [OGI-055], Innovative Medicines Initiative (EU/EFPIA) [ULTRA-DD grant no. 115766], Janssen, Merck & Co., Novartis Pharma AG, Ontario Ministry of Research, Innovation and Science (MRIS), Pfizer, São Paulo Research Foundation-FAPESP, Takeda, and the Wellcome Trust.

References

- (1) Sharff, A.; Jhoti, H. High-Throughput Crystallography to Enhance Drug Discovery. *Curr. Opin. Chem. Biol.* **2003**, *7*, 340–345.
- (2) Renaud, J.-P.; Chung, C.; Danielson, U. H.; Egner, U.; Hennig, M.; Hubbard, R. E.; Nar, H. Biophysics in Drug Discovery: Impact, Challenges and Opportunities. *Nat. Rev. Drug Discov.* **2016**, *15*, 679–698.
- (3) Todd, A. E.; Marsden, R. L.; Thornton, J. M.; Orengo, C. A. Progress of Structural Genomics Initiatives: An Analysis of Solved Target Structures. *J. Mol. Biol.* **2005**, *348*, 1235–1260.
- (4) Yee, A.; Pardee, K.; Christendat, D.; Savchenko, A.; Edwards, A. M.; Arrowsmith, C. H. Structural Proteomics: Toward High-Throughput Structural Biology as a Tool in Functional Genomics. *Acc. Chem. Res.* **2003**, *36*, 183–189.
- (5) Schapira, M.; Structural Genomics, Its Application in Chemistry, Biology, and Drug Discovery. In *Burger's Medicinal Chemistry and Drug Discovery*; John Wiley & Sons, Inc., 2003; pp. 569–600.
- (6) Berman, H. M. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242.
- (7) Wang, R.; Fang, X.; Lu, Y.; Wang, S. The PDBbind Database: Collection of Binding Affinities for Protein-Ligand Complexes with Known Three-Dimensional Structures. *J. Med. Chem.* **2004**, *47*, 2977–2980.
- (8) Wang, R.; Fang, X.; Lu, Y.; Yang, C. Y.; Wang, S. The PDBbind Database: Methodologies and Updates. *J. Med. Chem.* **2005**, *48*, 4111–4119.
- (9) Deng, Z.; Chuaqui, C.; Singh, J. Structural Interaction Fingerprint (SIFt): A Novel

- Method for Analyzing Three-Dimensional Protein-Ligand Binding Interactions. *J. Med. Chem.* **2004**, *47*, 337–344.
- (10) Verdonk, M. L.; Cole, J. C.; Taylor, R. SuperStar: A Knowledge-Based Approach for Identifying Interaction Sites in Proteins. *J. Mol. Biol.* **1999**, *289*, 1093–1108.
- (11) Hendlich, M.; Bergner, A.; Günther, J.; Klebe, G. Relibase: Design and Development of a Database for Comprehensive Analysis of Protein–Ligand Interactions. *J. Mol. Biol.* **2003**, *326*, 607–620.
- (12) Golovin, A.; Henrick, K. MSDmotif: Exploring Protein Sites and Motifs. *BMC Bioinformatics* **2008**, *9*, 1–11.
- (13) Inhester, T.; Bietz, S.; Hilbig, M.; Schmidt, R.; Rarey, M. Index-Based Searching of Interaction Patterns in Large Collections of Protein–Ligand Interfaces. *J. Chem. Inf. Model.* **2017**, *57*, 148–158.
- (14) Marcou, G.; Rognan, D. Optimizing Fragment and Scaffold Docking by Use of Molecular Interaction Fingerprints. *J. Chem. Inf. Model.* **2007**, *47*, 195–207.
- (15) Salentin, S.; Schreiber, S.; Haupt, V. J.; Adasme, M. F.; Schroeder, M. PLIP: Fully Automated Protein-Ligand Interaction Profiler. *Nucleic Acids Res.* **2015**, *43*, W443–W447.
- (16) Jubb, H. C.; Higuero, A. P.; Ochoa-Montano, B.; Pitt, W. R.; Ascher, D. B.; Blundell, T. L. Arpeggio: A Web Server for Calculating and Visualising Interatomic Interactions in Protein Structures. *J. Mol. Biol.* **2016**, *429*, 365–371.
- (17) Schreyer, A. M.; Blundell, T. L. CREDO: A Structural Interactomics Database for Drug Discovery. *Database* **2013**, *2013*, 1–9.
- (18) Mohamed, R.; Degac, J.; Helms, V. Composition of Overlapping Protein-Protein

- and Protein-Ligand Interfaces. *PLoS One* **2015**, *10*, 1–18.
- (19) Chen, K.; Kurgan, L. Investigation of Atomic Level Patterns in Protein-Small Ligand Interactions. *PLoS One* **2009**, *4*, 1–14.
- (20) Reynolds, C. H.; Tounge, B. A.; Bembenek, S. D. Ligand Binding Efficiency: Trends, Physical Basis, and Implications. *J. Med. Chem.* **2008**, *51*, 2432–2438.
- (21) Hann, M. M. Molecular Obesity, Potency and Other Addictions in Drug Discovery. *Multifaceted Roles Crystallogr. Mod. Drug Discov.* **2015**, *2*, 183–196.
- (22) Erlanson, D. A.; Fesik, S. W.; Hubbard, R. E.; Jahnke, W.; Jhoti, H. Twenty Years on: The Impact of Fragments on Drug Discovery. *Nat. Rev. Drug Discov.* **2016**, *15*, 605–619.
- (23) Smith, R. D.; Engdahl, A. L.; Dunbar, J. B.; Carlson, H. A. Biophysical Limits of Protein-Ligand Binding. *J. Chem. Inf. Model.* **2012**, *52*, 2098–2106.
- (24) Schaeffer, L. The Role of Functional Groups in Drug–Receptor Interactions. In *The Practice of Medicinal Chemistry*; Elsevier, 2008; pp. 464–480.
- (25) Higuieruelo, A. P.; Schreyer, A.; Bickerton, G. R. J.; Blundell, T. L.; Pitt, W. R. What Can We Learn from the Evolution of Protein-Ligand Interactions to Aid the Design of New Therapeutics? *PLoS One* **2012**, *7*, 1–8.
- (26) Ritchie, T. J.; Macdonald, S. J. F. Physicochemical Descriptors of Aromatic Character and Their Use in Drug Discovery. *J. Med. Chem.* **2014**, *57*, 7206–7215.
- (27) Taylor, R. D.; MacCoss, M.; Lawson, A. D. G. Rings in Drugs. *J. Med. Chem.* **2014**, *57*, 5845–5859.
- (28) Meyer, E. A.; Castellano, R. K.; Diederich, F. Interactions with Aromatic Rings in Chemical and Biological Recognition. *Angew. Chem. Int. Ed. Engl.* **2003**, *42*,

1210–1250.

- (29) Brandl, M.; Weiss, M. S.; Jabs, A.; Sühnel, J.; Hilgenfeld, R. C-H \cdots π -Interactions in Proteins. *J. Mol. Biol.* **2001**, *307*, 357–377.
- (30) Nishio, M.; Umezawa, Y.; Fantini, J.; Weiss, M. S.; Chakrabarti, P. CH- π Hydrogen Bonds in Biological Macromolecules. *Phys. Chem. Chem. Phys.* **2014**, *16*, 12648–12683.
- (31) Saraogi, I.; Vijay, V. G.; Das, S.; Sekar, K.; Guru Row, T. N. C–halogen \cdots π Interactions in Proteins: A Database Study. *Cryst. Eng.* **2003**, *6*, 69–77.
- (32) Imai, Y. N.; Inoue, Y.; Nakanishi, I.; Kitaura, K. Cl- π Interactions in Protein-Ligand Complexes. *Protein Sci.* **2008**, *17*, 1129–1137.
- (33) Valley, C. C.; Cembran, A.; Perlmutter, J. D.; Lewis, A. K.; Labello, N. P.; Gao, J.; Sachs, J. N. The Methionine-Aromatic Motif Plays a Unique Role in Stabilizing Protein Structure. *J. Biol. Chem.* **2012**, *287*, 34979–34991.
- (34) Davis, A. M.; Teague, S. J. Hydrogen Bonding, Hydrophobic Interactions, and Failure of the Rigid Receptor Hypothesis. *Angew. Chemie Int. Ed.* **1999**, *38*, 736–749.
- (35) Leung, C. S.; Leung, S. S. F.; Tirado-Rives, J.; Jorgensen, W. L. Methyl Effects on Protein-Ligand Binding. *J. Med. Chem.* **2012**, *55*, 4489–4500.
- (36) Hajduk, P. J.; Sauer, D. R. Statistical Analysis of the Effects of Common Chemical Substituents on Ligand Potency. *J. Med. Chem.* **2008**, *51*, 553–564.
- (37) de Vicente, J.; Tivitmahaisoon, P.; Berry, P.; Bolin, D. R.; Carvajal, D.; He, W.; Huang, K.-S.; Janson, C.; Liang, L.; Lukacs, C.; Petersen, A.; Qian, H.; Yi, L.; Zhuang, Y.; Hermann, J. C. Fragment-Based Drug Design of Novel

- Pyranopyridones as Cell Active and Orally Bioavailable Tankyrase Inhibitors. *ACS Med. Chem. Lett.* **2015**, *6*, 1019–1024.
- (38) Schönherr, H.; Cernak, T. Profound Methyl Effects in Drug Discovery and a Call for New C-H Methylation Reactions. *Angew. Chemie - Int. Ed.* **2013**, *52*, 12256–12267.
- (39) Bissantz, C.; Kuhn, B.; Stahl, M. A Medicinal Chemist's Guide to Molecular Interactions. *J. Med. Chem.* **2010**, *53*, 5061–5084.
- (40) Steiner, T. The Hydrogen Bond in the Solid State. *Angew. Chem. Int. Ed. Engl.* **2002**, *41*, 49–76.
- (41) Fersht, A. R. The Hydrogen Bond in Molecular Recognition. *Trends Biochem. Sci.* **1987**, *12*, 301–304.
- (42) Watson, J. D.; Crick, F. H. C. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature* **1953**, *171*, 737–738.
- (43) Pauling, L.; Corey, R. B.; Branson, H. R. The Structure of Proteins; Two Hydrogen-Bonded Helical Configurations of the Polypeptide Chain. *Proc. Natl. Acad. Sci. U. S. A.* **1951**, *37*, 205–211.
- (44) Baker, E. N.; Hubbard, R. E. Hydrogen Bonding in Globular Proteins. *Prog. Biophys. Mol. Biol.* **1984**, *44*, 97–179.
- (45) Fersht, a R.; Shi, J. P.; Knill-Jones, J.; Lowe, D. M.; Wilkinson, a J.; Blow, D. M.; Brick, P.; Carter, P.; Waye, M. M.; Winter, G. Hydrogen Bonding and Biological Specificity Analysed by Protein Engineering. *Nature* **1985**, *314*, 235–238.
- (46) Kuhn, B.; Mohr, P.; Stahl, M. Intramolecular Hydrogen Bonding in Medicinal

- Chemistry. *J. Med. Chem.* **2010**, *53*, 2601–2611.
- (47) H. Williams, D.; S. Westwell, M.; Pawlak, K.; Bruening, R. L.; Tarbet, B. J.; Marshall, G. R.; Brick, P.; Carter, P.; Waye, M. M. Y.; Winter, G. Aspects of Weak Interactions. *Chem. Soc. Rev.* **1998**, *27*, 57–64.
- (48) Shoichet, B. K. No Free Energy Lunch. *Nat. Biotechnol.* **2007**, *25*, 1109–1110.
- (49) Freire, E. Do Enthalpy and Entropy Distinguish First in Class from Best in Class? *Drug Discov. Today* **2008**, *13*, 869–874.
- (50) Muley, L.; Baum, B.; Smolinski, M.; Freindorf, M.; Heine, A.; Klebe, G.; Hangauer, D. G. Enhancement of Hydrophobic Interactions and Hydrogen Bond Strength by Cooperativity: Synthesis, Modeling, and Molecular Dynamics Simulations of a Congeneric Series of Thrombin Inhibitors. *J. Med. Chem.* **2010**, *53*, 2126–2135.
- (51) Baum, B.; Muley, L.; Heine, A.; Smolinski, M.; Hangauer, D.; Klebe, G. Think Twice: Understanding the High Potency of Bis(phenyl)methane Inhibitors of Thrombin. *J. Mol. Biol.* **2009**, *391*, 552–564.
- (52) Tsuzuki, S.; Honda, K.; Uchimaru, T.; Mikami, M.; Tanabe, K. Origin of Attraction and Directionality of the Π/π Interaction: Model Chemistry Calculations of Benzene Dimer Interaction. *J. Am. Chem. Soc.* **2002**, *124*, 104–112.
- (53) Salonen, L. M.; Ellermann, M.; Diederich, F. Aromatic Rings in Chemical and Biological Recognition: Energetics and Structures. *Angew. Chemie - Int. Ed.* **2011**, *50*, 4808–4842.
- (54) Eldrup, A. B.; Soleymanzadeh, F.; Taylor, S. J.; Muegge, I.; Farrow, N. A.; Joseph, D.; McKellop, K.; Man, C. C.; Kukulka, A.; De Lombaert, S. Structure-

- Based Optimization of Arylamides as Inhibitors of Soluble Epoxide Hydrolase. *J. Med. Chem.* **2009**, *52*, 5880–5895.
- (55) Ritchie, T. J.; Macdonald, S. J. F. The Impact of Aromatic Ring Count on Compound Developability – Are Too Many Aromatic Rings a Liability in Drug Design? *Drug Discov. Today* **2009**, *14*, 1011–1020.
- (56) Lovering, F.; Bikker, J.; Humblet, C. Escape from Flatland: Increasing Saturation as an Approach to Improving Clinical Success. *J. Med. Chem.* **2009**, *52*, 6752–6756.
- (57) Steiner, T. Donor and Acceptor Strengths in C-H···O Hydrogen Bonds Quantified from Crystallographic Data of Small Solvent Molecules. *New J. Chem.* **1998**, *22*, 1099–1103.
- (58) Scheiner, S. Weak H-Bonds. Comparisons of CH···O to NH···O in Proteins and PH···N to Direct P···N Interactions. *Phys. Chem. Chem. Phys.* **2011**, *13*, 13860–13872.
- (59) Panigrahi, S. K.; Desiraju, G. R. Strong and Weak Hydrogen Bonds in the Protein-Ligand Interface. *Proteins Struct. Funct. Genet.* **2007**, *67*, 128–141.
- (60) Sarkhel, S.; Desiraju, G. R. N-H···O, O-H···O, and C-H···O Hydrogen Bonds in Protein-Ligand Complexes: Strong and Weak Interactions in Molecular Recognition. *Proteins Struct. Funct. Bioinforma.* **2003**, *54*, 247–259.
- (61) Derewenda, Z. S.; Lee, L.; Derewenda, U. The Occurrence of C-H···O Hydrogen Bonds in Proteins. *J. Mol. Biol.* **1995**, *252*, 248–262.
- (62) Wahl, M. C-H···O Hydrogen Bonding in Biology. *Trends Biochem. Sci.* **1997**, *22*, 97–102.

- (63) Jiang, L.; Lai, L. CH \cdots O Hydrogen Bonds at Protein-Protein Interfaces. *J. Biol. Chem.* **2002**, *277*, 37732–37740.
- (64) Vargas, R.; Garza, J.; Dixon, D. A.; Hay, B. P. How Strong Is the C α –H \cdots OC Hydrogen Bond? *J. Am. Chem. Soc.* **2000**, *122*, 4750–4755.
- (65) Liu, Z.; Wang, G.; Li, Z.; Wang, R. Geometrical Preferences of the Hydrogen Bonds on Protein Ligand Binding Interface Derived from Statistical Surveys and Quantum Mechanics Calculations. *J. Chem. Theory Comput.* **2008**, *4*, 1959–1973.
- (66) Keegstra, E. M. D.; Spek, A. L.; Zwikker, J. W.; Jenneskens, L. W. The Crystal Structure of 2-Methoxy-1,4-Benzoquinone: Molecular Recognition Involving Intermolecular Dipole-Dipole and C-H \cdots O Hydrogen Bond Interactions. *J. Chem. Soc. Chem. Commun.* **1994**, *7*, 1633–1634.
- (67) Aravinda, S.; Shamala, N.; Bandyopadhyay, A.; Balaram, P. Probing the Role of the C-H \cdots O Hydrogen Bond Stabilized Polypeptide Chain Reversal at the C-Terminus of Designed Peptide Helices. Structural Characterization of Three Decapeptides. *J. Am. Chem. Soc.* **2003**, *125*, 15065–15075.
- (68) Mandel-Gutfreund, Y.; Margalit, H.; Jernigan, R. L.; Zhurkin, V. B. A Role for CH \cdots O Interactions in Protein-DNA Recognition. *J Mol Biol* **1998**, *277*, 1129–1140.
- (69) Horowitz, S.; Trievel, R. C. Carbon-Oxygen Hydrogen Bonding in Biological Structure and Function. *J. Biol. Chem.* **2012**, *287*, 41576–41582.
- (70) Musah, R. A.; Jensen, G. M.; Rosenfeld, R. J.; McRee, D. E.; Goodin, D. B. Variation in Strength of an Unconventional C-H \cdots O Hydrogen Bond in an Engineered Protein Cavity. *J. Am. Chem. Soc.* **1997**, *119*, 9083–9084.

- (71) Pierce, A. C.; Sandretto, K. L.; Bemis, G. W. Kinase Inhibitors and the Case for CH \cdots O Hydrogen Bonds in Protein-Ligand Binding. *Proteins Struct. Funct. Genet.* **2002**, *49*, 567–576.
- (72) Wang, S.; Meades, C.; Wood, G.; Osnowski, A.; Anderson, S.; Yuill, R.; Thomas, M.; Mezna, M.; Jackson, W.; Midgley, C.; Griffiths, G.; Fleming, I.; Green, S.; McNae, I.; Wu, S. Y.; McInnes, C.; Zheleva, D.; Walkinshaw, M. D.; Fischer, P. M. 2-Anilino-4-(Thiazol-5-Yl)pyrimidine CDK Inhibitors: Synthesis, SAR Analysis, X-Ray Crystallography, and Biological Activity. *J. Med. Chem.* **2004**, *47*, 1662–1675.
- (73) UniProtKB/TrEMBL 2017_03 <http://www.uniprot.org/statistics/TrEMBL> (accessed Mar 28, 2017).
- (74) Nagy, P. I.; Erhardt, P. W. Theoretical Studies of Salt-Bridge Formation by Amino Acid Side Chains in Low and Medium Polarity Environments. *J. Phys. Chem. B* **2010**, *114*, 16436–16442.
- (75) Waldburger, C. D.; Schildbach, J. F.; Sauer, R. T. Are Buried Salt Bridges Important for Protein Stability and Conformational Specificity? *Nat. Struct. Biol.* **1995**, *2*, 122–128.
- (76) Hendsch, Z. S.; Tidor, B. Do Salt Bridges Stabilize Proteins - a Continuum Electrostatic Analysis. *Prot. Sci.* **1994**, *3*, 211–226.
- (77) Segala, E.; Guo, D.; Cheng, R. K. Y.; Bortolato, A.; Deflorian, F.; Doré, A. S.; Errey, J. C.; Heitman, L. H.; Ijzerman, A. P.; Marshall, F. H.; Cooke, R. M. Controlling the Dissociation of Ligands from the Adenosine A2A Receptor through Modulation of Salt Bridge Strength. *J. Med. Chem.* **2016**, *59*, 6470–6479.

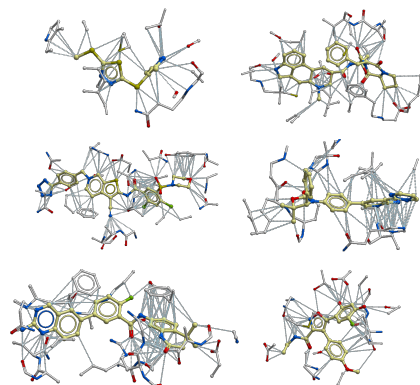
- (78) Miyamoto, Y.; Banno, Y.; Yamashita, T.; Fujimoto, T.; Oi, S.; Moritoh, Y.; Asakawa, T.; Kataoka, O.; Yashiro, H.; Takeuchi, K.; Suzuki, N.; Ikedo, K.; Kosaka, T.; Tsubotani, S.; Tani, A.; Sasaki, M.; Funami, M.; Amano, M.; Yamamoto, Y.; Aertgeerts, K.; Yano, J.; Maezaki, H. Discovery of a 3-Pyridylacetic Acid Derivative (TAK-100) as a Potent, Selective and Orally Active Dipeptidyl Peptidase IV (DPP-4) Inhibitor. *J. Med. Chem.* **2011**, *54*, 831–850.
- (79) Weltrowska, G.; Chung, N. N.; Lemieux, C.; Guo, J.; Lu, Y.; Wilkes, B. C.; Schiller, P. W. “Carba”-Analogues of Fentanyl Are Opioid Receptor Agonists. *J. Med. Chem.* **2010**, *53*, 2875–2881.
- (80) Peng, Y. H.; Shiao, H. Y.; Tu, C. H.; Liu, P. M.; Hsu, J. T. A.; Amancha, P. K.; Wu, J. S.; Coumar, M. S.; Chen, C. H.; Wang, S. Y.; Lin, W. H.; Sun, H. Y.; Chao, Y. S.; Lyu, P. C.; Hsieh, H. P.; Wu, S. Y. Protein Kinase Inhibitor Design by Targeting the Asp-Phe-Gly (DFG) Motif: The Role of the DFG Motif in the Design of Epidermal Growth Factor Receptor Inhibitors. *J. Med. Chem.* **2013**, *56*, 3889–3903.
- (81) Harder, M.; Kuhn, B.; Diederich, F. Efficient Stacking on Protein Amide Fragments. *ChemMedChem* **2013**, *8*, 397–404.
- (82) Mitchell, J. B.; Nandi, C. L.; McDonald, I. K.; Thornton, J. M.; Price, S. L. Amino/aromatic Interactions in Proteins: Is the Evidence Stacked against Hydrogen Bonding? *J. Mol. Biol.* **1994**, *239*, 315–331.
- (83) Duan, G.; Smith, V. H.; Weaver, D. F. An Ab Initio and Data Mining Study on Aromatic–amide Interactions. *Chem. Phys. Lett.* **1999**, *310*, 323–332.
- (84) Giroud, M.; Ivkovic, J.; Martignoni, M.; Fleuti, M.; Trapp, N.; Haap, W.;

- Kuglstatter, A.; Benz, J.; Kuhn, B.; Schirmeister, T.; Diederich, F. Inhibition of the Cysteine Protease Human Cathepsin L by Triazine Nitriles: Amide···Heteroarene π -Stacking Interactions and Chalcogen Bonding in the S3 Pocket. *ChemMedChem* **2016**, *12*, 257–270.
- (85) Giroud, M.; Harder, M.; Kuhn, B.; Haap, W.; Trapp, N.; Schweizer, W. B.; Schirmeister, T.; Diederich, F. Fluorine Scan of Inhibitors of the Cysteine Protease Human Cathepsin L: Dipolar and Quadrupolar Effects in the π -Stacking of Fluorinated Phenyl Rings on Peptide Amide Bonds. *ChemMedChem* **2016**, *11*, 1042–1047.
- (86) Qiu, Z.; Kuhn, B.; Aebi, J.; Lin, X.; Ding, H.; Zhou, Z.; Xu, Z.; Xu, D.; Han, L.; Liu, C.; Qiu, H.; Zhang, Y.; Haap, W.; Riemer, C.; Stahl, M.; Qin, N.; Shen, H. C.; Tang, G. Discovery of Fluoromethylketone-Based Peptidomimetics as Covalent ATG4B (Autophagin-1) Inhibitors. *ACS Med. Chem. Lett.* **2016**, *7*, 802–806.
- (87) Levin, E. J.; Quick, M.; Zhou, M. Crystal Structure of a Bacterial Homologue of the Kidney Urea Transporter. *Nature* **2009**, *462*, 757–761.
- (88) Salonen, L. M.; Holland, M. C.; Kaib, P. S. J.; Haap, W.; Benz, J.; Mary, J.-L.; Kuster, O.; Schweizer, W. B.; Banner, D. W.; Diederich, F. Molecular Recognition at the Active Site of Factor Xa: Cation– π Interactions, Stacking on Planar Peptide Surfaces, and Replacement of Structural Water. *Chem. – A Eur. J.* **2012**, *18*, 213–222.
- (89) Mahadevi, A. S.; Sastry, G. N. Cation– π Interaction: Its Role and Relevance in Chemistry, Biology, and Material Science. *Chem. Rev.* **2013**, *113*, 2100–2138.
- (90) Gallivan, J. P.; Dougherty, D. A. Cation-Pi Interactions in Structural Biology.

Proc. Natl. Acad. Sci. U. S. A. **1999**, *96*, 9459–9464.

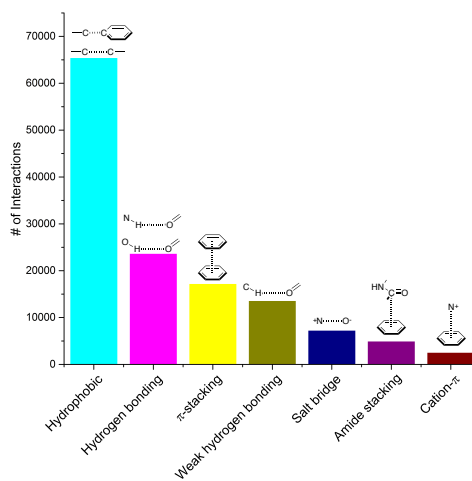
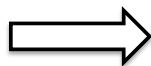
- (91) Dougherty, D. A. The Cation- π Interaction. *Acc. Chem. Res.* **2013**, *46*, 885–893.
- (92) Kamps, J. J. a. G.; Huang, J.; Poater, J.; Xu, C.; Pieters, B. J. G. E.; Dong, A.; Min, J.; Sherman, W.; Beuming, T.; Matthias Bickelhaupt, F.; Li, H.; Mecinović, J. Chemical Basis for the Recognition of Trimethyllysine by Epigenetic Reader Proteins. *Nat. Commun.* **2015**, *6*, 1–12.
- (93) Lummis, S. C. R.; Harrison, N. J.; Wang, J.; Ashby, J. A.; Millen, K. S.; Beene, D. L.; Dougherty, D. A. Multiple Tyrosine Residues Contribute to GABA Binding in the GABA C Receptor Binding Pocket. *ACS Chem. Neurosci.* **2012**, *3*, 186–192.
- (94) Pless, S. a; Hanek, A. P.; Price, K. L.; Lynch, J. W.; Lester, H. a; Dougherty, D. a; Lummis, S. C. R. A Cation- π Interaction at a Phenylalanine Residue in the Glycine Receptor Binding Site Is Conserved for Different Agonists. *Mol. Pharmacol.* **2011**, *79*, 742–748.
- (95) Duffy, N. H.; Lester, H. A.; Dougherty, D. A. Ondansetron and Granisetron Binding Orientation in the 5-HT 3 Receptor Determined by Unnatural Amino Acid Mutagenesis. *ACS Chem. Biol.* **2012**, *7*, 1738–1745.
- (96) Salonen, L. M.; Bucher, C.; Banner, D. W.; Haap, W.; Mary, J. L.; Benz, J.; Kuster, O.; Seiler, P.; Schweizer, W. B.; Diederich, F. Cation- π Interactions at the Active Site of Factor Xa: Dramatic Enhancement upon Stepwise N-Alkylation of Ammonium Ions. *Angew. Chemie - Int. Ed.* **2009**, *48*, 811–814.

Table of Contents Graphic



PDB

11,016 protein-ligand complexes
750,873 protein-ligand atomic interactions



Top interactions for structure-based drug design