Journal of Molecular Graphics and Modelling 27 (2008) 170–177

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

Halogenated ligands and their interactions with amino acids: Implications for structure–activity and structure–toxicity relationships

Sandhya Kortagere [a], Sean Ekins [a,b,c], William J. Welsh [a,*]

[a] Department of Pharmacology and Environmental Bioinformatics & Computational Toxicology Center (ebCTC), University of Medicine & Dentistry of New Jersey (UMDNJ)-Robert Wood Johnson Medical School, 675 Hoes lane, Piscataway, NJ 08854, USA
[b] Collaborations in Chemistry, 601 Runnymede Avenue, Jenkintown, PA 19046, USA
[c] Department of Pharmaceutical Sciences, University of Maryland, MD 21201, USA

ARTICLE INFO

Article history:
Received 17 December 2007
Received in revised form 2 April 2008
Accepted 3 April 2008
Available online 10 April 2008

Keywords:
DDT
Halogenated ligands
Halogen bonds
PCB
PDB analysis
Propensity
Protein–ligand interactions
Structure–activity relationship and structure–toxicity relationship

ABSTRACT

The properties of chemicals are rooted in their molecular structure. It follows that structural analysis of specific interactions between ligands and biomolecules at the molecular level is invaluable for defining structure–activity relationships (SARs) and structure–toxicity relationships (STRs). This study has elucidated the structural and molecular basis of interactions of biomolecules with alkyl and aryl halides that are extensively used as components in many commercial pesticides, disinfectants, and drugs. We analyzed the protein structures deposited in Protein Data Bank (PDB) for structural information associated with interactions between halogenated ligands and proteins. This analysis revealed distinct patterns with respect to the nature and structural characteristics of halogen interactions with specific types of atoms and groups in proteins. Fluorine had the highest propensity of interactions for glycine, while chlorine for leucine, bromine for arginine, and iodine for lysine. Chlorine, bromine and iodine had the lowest propensity of interactions for cysteine, while fluorine had a lowest propensity for proline. These trends for highest propensity shifted towards the hydrophobic residues for all the halogens when only interactions with the side chain were considered. Halogens had equal propensities of interaction for the halogen bonding partners (nitrogen and oxygen atoms), albeit with different geometries. The optimal angle for interactions with halogens was ~120° for oxygen atoms, and ~96° for nitrogen atoms. The distance distributions of halogens with various amino acids were mostly bimodal, and the angle distributions were unimodal. Insights gained from this study have implications for the rational design of safer drugs and commercially important chemicals.

1. Introduction

Halogenation of pharmacologically active compounds has long been a common method to increase their metabolic stability and lipophilicity; however, their use can sometimes lead to undesirable effects including toxicity. Often, the metabolic stability afforded by halogenation of a lead compound leads to increased binding affinity to the receptor, channel, transporter or other protein target. The utility of halogenation can be witnessed in numerous cases, of which due to space limitations we will mention only a small number. For example, fluorination of 3-(3-(piperidin-1-yl)propyl)indole and some piperazinyl analogues not only enhanced their oral absorption but also increased the binding affinity for the 5-HT_{1D} receptor [1]. Similarly, fluorination of 3-(4-piperidin-3-yl)-2-phenylindoles at the 6th position on the indole ring blocked oxidation of the compound, thereby simultaneously increasing its bioavailability and binding affinity for the 5-HT_{2A} receptor [2]. However, the increased affinity leads to toxic effects if this compound binds to an unintended, off- or anti-target [3]. Similar effects have been observed for many other compounds such as ion channel blockers [4], cholesterol inhibitors [5], antibiotics [6], and anti-cancer agents [7]. Most of the neuroleptic drugs in current use are halogenated at key positions, and the halogen atom forms an important component of their pharmacophore. Fluorine is the most frequently used halogen substitutent, but other halogens are used to enhance some desirable property (e.g., metabolic stability, shelf life, intestinal and/or blood–brain barrier permeability). For instance, many antibiotics such as vancomycin and cefaclor are chlorinated, drugs used to treat respiratory disorders contain bromine [8], and compounds used to treat hormonal imbalance and some forms of cancer commonly feature iodine substitution [9].

* Corresponding author. Tel.: +1 732 235 3234; fax: +1 732 235 3475.
E-mail address: welshwj@umdnj.edu (W.J. Welsh).

1093-3263/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2008.04.001

Halogens have also found applications in molecular imaging and structural studies. For example, halogenation of certain estradiol analogues increased the affinity for estrogen receptors in the order of Br ≫ Cl > I, while ^{123}I labeled estradiol has been used for SPECT imaging in the diagnosis of breast cancer [9]. A study by Gee et al. [10] showed that mono-halogenated derivatives of fura-2 and indo-1 could be used as probes to detect variations in intracellular calcium levels and the effectiveness of the probe was directly linked to the position of the halogen. Many halogenated compounds, such as bimane and 3-fluorotyrosyl green, are commonly used as fluorescent probes. Halogenation has also been used to modify the conformational properties of peptides; examples include the substitution of leucine by trifluoroleucine [11] or hexafluoroleucine [12] in the hydrophobic surface of a α-helical coiled coil peptide, which results in enhanced thermal as well as structural stability. Mono-halogenated aromatic derivatives of phenol and tyrosine have been used as fluorescent probes to study various aspects of protein structure such as ligand-induced conformational changes or the nature of transition states [13].

All the above examples illustrate that halogens are unique chemical elements that are widely employed in commercially available molecules to augment existing properties and, in some cases, to introduce new properties. In view of the dramatic impact of halogenation on the structure and properties of chemicals, it is imperative to thoroughly understand the molecular interactions between halogens and various amino acids and nucleic acids that form the building blocks of biological systems. Previous studies of halogen interactions among small molecules [14–17] have shown the importance of short-range electrostatics, reflecting the polar nature of the halogen atom engaged in charge-transfer interactions and hydrogen bond formation with other donor systems. In a recent survey [18], short-range interactions or "halogen bonds" between halogen atoms and backbone carbonyl atoms of proteins and various nuclei acids were determined to be functionally significant. However, these studies were conducted using a small data set and did not explore the nature or geometry of electrostatic and hydrophobic interactions.

In this study, we provide a detailed structural analysis of the interactions between halogens and the side-chain atoms of various amino acids from high-resolution protein–ligand complexes deposited in the Protein Data Bank (PDB) [19]. The purpose of the present study was to understand contributions by both the electrostatic (polar) and hydrophobic (nonpolar) nature of halogens in a protein environment. Specifically, the interaction distance and spatial orientations were analyzed and the propensities for interactions between each halogen type and the amino acids were estimated. We explored the alkyl- and aryl-substituted halogen compounds separately to determine the differences in their spatial orientation relative to the amino acids in terms of interplanar angles and hydrogen bonding interactions. By virtue of the extensive use of halogenated chemicals among pesticides, fungicides, pharmaceuticals, and other commercial products, we also investigated the halogenated ligand interactions of toxic chemicals.

2. Methods

The PDB (July 2006 release, www.rcsb.org/pdb) was queried for entries of halogenated ligands co-crystallized with proteins in order to map the interactions of halogen atoms with proteins. Only structures with 3 Å or better resolution were considered, resulting in a set of 3833 unique entries. All interactions between the halogenated atom of the ligand and any of the amino acid side-chain atoms less than 6 Å in distance were included. Only the closest distance to a non-hydrogen atom of an amino acid side chain was recorded as an interaction. Based on these criteria, the interacting atom pairs were culled using the neighbor module of HBPLUS (http://www.biochem.ucl.ac.uk/bsm/hbplus/home.html) [20] and filtered using a custom-made script. The propensity P of interaction of a halogen X (where X = F, Cl, Br or I) with each of the 20 naturally occurring amino acids in the entire data set was recorded as the observed frequency (OBS-Freq). Also, the expected frequency (EXP-Freq) was determined for each such interacting pair. The corresponding equations employed are given as

$$\text{OBS-Freq} = \frac{P(\text{A},\text{X})}{N(\text{X},\text{X})} \quad \text{and} \quad \text{EXP-Freq} = \frac{P(\text{A},\text{X})}{N(\text{A},\text{A})},$$

where P(A,X) is the probability of interaction of halogen X with amino acid A; N(X,X) is the total number of halogen interactions of a particular halogen atom X; and N(A,A) is the total number of amino acid interactions of a particular amino acid A in the data set. Finally, in order to compare the propensities of interactions for halogens with various amino acids, the normalized frequency (NOR-Freq) was calculated as

$$\text{NOR-freq} = \text{OBS-Freq} \times \text{EXP-Freq}.$$

We investigated the interaction geometry of halogenated ligands with proteins. An upper distance limit of 6 Å was used in analyzing the non-bonded interactions. Interactions within this distance included both hydrogen bonds and van der Waals types. In all cases, we analyzed the distance "*d*" and the corresponding angles formed by the halogen atom with the atoms of the amino acid side-chain represented as '*θ*' in Fig. 1. For all interactions the interplanar angle '*φ*' between the planes containing the interacting halogen system and the amino acid side chain was calculated (Fig. 1). In order to assess the nature of the interaction between halogen atoms and possible hydrogen bond donors (known as "halogen bonds"), we classified the amino acids into two separate classes: nitrogen containing (Arg-NH, Asn-ND, Gln-NE, His-ND, Lys-NZ, Trp-NE) and oxygen containing (Asp-OD, Glu-OE, Ser-OG, Thr-OG, Tyr-OH). Hydrogen bond interactions were limited to those within the sum of the van der Waals radii of the interacting atoms as described previously [18], with an allowance for coordinate correction. This resulted in an upper distance limit of 3.6 Å for oxygen and nitrogen mediated interactions. We also included the interactions of the halogenated ligands with crystallographic water molecules (using the same distance cut-off for

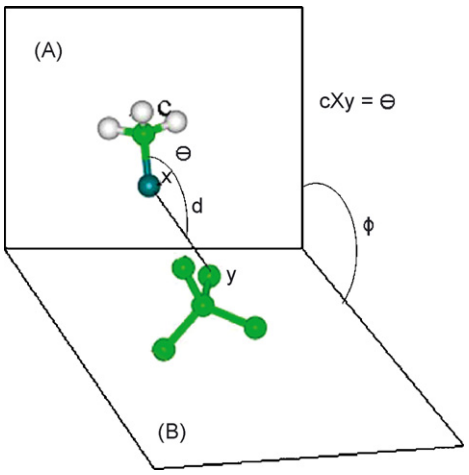

Fig. 1. Schematic representation of halogenated ligand interaction, where the 'X' represents the halogen atom (X = F, Cl, Br or I); *d* is the distance of interaction; *θ* is the angle formed by atoms c, X and y; and *φ* is the interplanar angle between planes A and B.

oxygen). No limitations were imposed on the angles, since our aim was to understand all potential interactions between the hydrogen bonding partners.

3. Results

The July 2006 release of the PDB contained 3833 entries with one or more halogenated ligand co-crystallized with a protein. Among these 3833 PDB entries, 186 proteins had co-crystallized ligands with more than one type of halogen atom substitution. Further, 657 of these entries contained only fluorinated ligands, 2526 had chlorinated, 211 had brominated and 253 had iodinated ligands. Since the data set contained more fluorinated and chlorinated ligands than brominated or iodinated ligands, the observed frequencies were normalized to allow for a statistically meaningful comparison between all the halogen classes. Also, the observed frequencies were normalized based on the occurrence of a particular amino acid in the binding site to yield the expected frequencies. Finally to compare the statistics of interactions of various halogens, the normalized frequency (NOR-Freq) was calculated. The observed frequency (OBS-Freq), expected frequency (EXP-Freq), and NOR-Freq of interactions of halogens with the 20 amino acids are summarized in Table 1.

Among all the interactions, fluorine had the highest propensity for interaction with glycine, chlorine with leucine, bromine with arginine and iodine with lysine. Conversely, fluorine had the lowest propensity of interaction with proline, while chlorine, bromine and iodine exhibited the lowest frequency of interactions with cysteine. However, there was a shift in this trend if only the side-chain interactions were considered. Fluorine and chlorine substitutions were found to have the highest propensities for interactions with the side-chain atoms of leucine, bromine with phenylalanine, and iodine with serine. Fluorine had the lowest frequency of interaction for the side chains of cysteine, chlorine for glutamine, bromine for aspartic acid, and iodine for proline.

In order to understand the behavior of halogens in the presence of charged groups, the interactions between the four halogens and the hydrogen bonding partners, viz., oxygen and nitrogen atoms of the 20 amino acids, were evaluated (interactions with water were analyzed separately for all the halogens). The number of occurrences for each of these interactions is shown in Table 2. Based on the normalized propensity values (Table 2), the halogen atoms have an equal propensity of forming hydrogen bonded interactions with both nitrogen and oxygen atoms. However, the geometry of these interactions varied between the two sets. The mean distance of interaction between the halogens and the hydrogen bonding partners increased steadily from fluorine to iodine in accordance with their electronegativity and radius profiles. The geometrical parameters of interaction (distance and angle) with oxygen and nitrogen atoms for the four halogens are listed in Table 2. A peak angle value was observed for the halogen interactions at ∼120° with oxygen atoms and at ∼96° with nitrogen atoms.

Further analysis was performed to clearly distinguish between the nature of interaction of halogen with C=O atoms. These interactions were classified into two categories: those with the main chain carbonyl and those with one of the side-chain carboxylate atoms of acidic amino acids. These interactions revealed certain distinct geometric variations between the two categories. The interactions of halogen atoms with the acidic side-chain oxygen of glutamic acid and aspartic acid had an peak angle at ∼120°. However, the interactions of halogens with the peptide

Table 1
Propensity of interaction between the various classes of halogen atoms of ligands and the amino acids in binding sites

AA	No. of interactions	F-OBS-Freq	F-EXP-Freq	F-NOR-Freq	No. of interactions	Cl-OBS-Freq	Cl-EXP-Freq	Cl-NOR-Freq	No. of interactions	Br-OBS-Freq	Br-EXP-Freq	Br-NOR-Freq	No. of interactions	I-OBS-Freq	I-EXP-Freq	I-NOR-Freq
ALA	1439	.05	.19	1.01	5055	.07	.66	4.49	385	.04	.05	0.21	726	.06	.10	0.62
ARG	1401	.05	.17	0.90	4621	.06	.57	3.53	1258	.14	.16	***2.11***	801	.07	.10	0.71
ASN	954	.03	.20	0.71	2926	.04	.62	2.42	360	.04	.08	0.29	500	.04	.11	0.48
ASP	1715	.06	.36	2.29	2276	.02	.48	1.45	343	.04	.07	0.26	456	.04	.10	0.39
CYS	613	.02	.30	0.69	1204	.03	.59	***0.95***	83	.01	.04	***0.04***	133	.01	.07	***0.08***
GLN	790	.02	.20	0.58	2279	.03	.57	1.73	502	.05	.13	0.68	433	.04	.11	0.42
GLU	1071	.04	.25	0.99	2135	.06	.49	1.40	567	.06	.13	0.79	559	.05	.13	0.64
GLY	2631	.10	.36	***3.09***	4668	.07	.56	3.49	499	.05	.06	0.32	535	.05	.06	0.31
HIS	1133	.04	.17	0.70	4931	.06	.73	4.78	339	.04	.05	0.18	395	.04	.06	0.20
ILE	1421	.05	.19	1.03	4765	.12	.65	4.16	513	.06	.07	0.39	593	.05	.08	0.43
LEU	2392	.09	.18	1.56	9254	.05	.68	***8.39***	862	.09	.06	0.59	1144	.10	.08	0.85
LYS	1295	.05	.21	1.02	3389	.03	.55	2.51	520	.06	.09	0.48	906	.08	.15	***1.19***
MET	924	.03	.27	0.91	2120	.07	.61	1.72	215	.02	.06	0.14	225	.02	.06	0.13
PHE	1660	.06	.21	1.28	5159	.03	.64	4.44	552	.06	.07	0.41	641	.06	.08	0.46
PRO	680	.03	.17	***0.43***	2366	.06	.59	1.87	362	.04	.09	0.35	595	.05	.15	0.79
SER	1679	.06	.22	1.39	4658	.05	.62	3.85	597	.06	.08	0.51	603	.05	.08	0.43
THR	1666	.06	.25	1.57	3952	.02	.60	3.18	382	.04	.06	0.24	570	.05	.09	0.44
TRP	678	.03	.24	0.60	1766	.02	.62	1.47	139	.01	.05	0.07	258	.02	.09	0.21
TYR	1244	.05	.21	0.97	3814	.05	.64	3.26	414	.04	.07	0.31	488	.04	.08	0.35
VAL	1479	.06	.25	1.36	3439	.05	.58	2.65	392	.04	.07	0.28	667	.06	.11	0.66

The total number of interactions for each halogen, OBS-Freq, EXP-Freq and the NOR-Freq are listed (see Section 2 for further details). The highest and lowest propensities are indicated in boldfaced italics.

Table 2
Interaction parameters of halogenated ligands with hydrogen bonding partners are shown

Property	F···N	F···O	Cl···N	Cl···O	Br···N	Br···O	I···N	I···O
No. of interactions	6862	7615	21394	18358	2827	2289	2560	2704
NOR-Freq	25.54	28.35	28.61	24.55	30.45	24.65	23.60	24.08
Mean distance	3.05	2.8	3.1	2.9	3.2	3.05	3.3	3.15
Mean angle	96.09	120.55	92.69	120.04	101.23	118.91	95.88	117.96

The mean distance is measured in Angstroms, while the angles are measured in degrees. NOR-Freq represents the normalized propensity of interactions (see Section 2).

backbone carbonyl oxygen peaked at ∼100°. We culled all the interactions of halogens that could potentially form hydrogen bonds with the oxygen atom of water molecules, and found that in all cases the distance distributions were unimodal with a peak value of 3.4 Å and peak angle at 130°.

A detailed analysis of the nature of the interaction such as distance, angle and interplanar angles between halogens and all the amino acids was performed. The distances and corresponding angles of interaction for individual amino acids with fluorine atoms are illustrated in Fig. 2. The distance distributions were mostly bimodal (Fig. 2A), while the angle distributions were found to be unimodal (Fig. 2B). Similar distributions for distance and angles were observed for other halogens. Both the minimum and mean distances for the entire distribution of interactions between the halogens and amino acids are shown in supplementary Table 1. As the distance distributions were skewed, the skewness of the distribution was calculated. The distribution of interplanar angles formed between the aryl-X (where X = F, Cl, Br or I) ligands and the aromatic amino acids is shown in Fig. 3. The interplanar peak angle for interactions with aromatic amino acids is ∼90° for aryl-F and aryl-Cl, and ∼180° (or ∼0°) for aryl-Br and aryl-I.

4. Discussion

Non-covalent interactions involving halogens are important considerations in the formulation of meaningful structure–activity relationships (SARs) and structure–toxicity relationships (STRs). Our study summarizes key features of these interactions from structurally well-defined ligand–protein complexes expressed as the normalized propensity of interaction between the halogen and the individual amino acid residues. Based on the correlation analysis of observed and expected frequencies with the normalized frequencies, we found that halogen interactions with amino acids are not driven by the size of the amino acids. The propensity evaluations indicate that halogens have high propensity to interact both with hydrophobic amino acids such as leucine and phenylalanine and with the hydrogen-bonding residues serine or threonine. This feature of halogens offers utility for drug design. For example, in a study on the binding of fluorinated inhibitors to the thrombin receptor [21], the efficacy of the ligand to act as an inhibitor was attributed to the polar interactions of the fluorine atom with the backbone carbonyl atom of the protein and also to the edge-to-face interaction of the aryl-F ring with the trptophan's indole ring. Incorporating a halogen atom into the pharmacophore of the inhibitor led to the design of a number of high affinity ligands for the thrombin receptor [22].

Halogen atoms are known to form so-called halogen bonds with nitrogen and oxygen atoms. Although similar in some ways to hydrogen bonds, halogen bonds are much weaker in strength. However, the presence of halogen bonds between the halogenated ligand and donor atoms significantly alters its affinity to the target protein receptor [23]. In fact, there is strong evidence that α-haloesters, when placed in close proximity to carbonyl groups, exhibit greater toxicity due to the increased "leaving" ability of the halogen atom [24]. This study also confirmed that increasing the number of halogen atoms altered the mechanism of toxic action from being baseline narcosis to that of direct electrophilicity [24,25], which concurs with our present results. Specifically, we found that the interaction parameters such as mean distance and angle of interaction between the halogen with nitrogen atoms differed significantly from that with oxygen atoms (Table 2). The mean distance varied from 3.05 to 3.30 Å in going from fluorine to iodine, in accordance with their atomic size, while no such pattern was evident in the mean angle values. The example illustrated here is for polychlorinated biphenyl (PCB), where its chlorine atom forms a hydrogen bond with the backbone nitrogen atom of isoleucine in the binding site of human estrogen sulfotransferase [26] (Fig. 4A).

The interaction angle of halogens with the backbone oxygen atoms was found to be ∼120°. The interaction peak angle for fluorinated ligands interacting with oxygen is consistent with the results obtained in the study on halogen interactions with backbone carbonyls [18], where the peak angle of interaction was found to be ∼113°. However, the interactions of halogen atoms with oxygen atoms from amino acid side chains showed several unique features. For example, the distance and angle distributions between fluorinated ligands and the oxygen atoms from acidic side chains (OA type) were distinct from the distributions of the oxygen atoms from other side chains (OL-type). The distance distribution for F-OA type of interactions was bimodal while the F-OL type was unimodal. This difference may be due to the additional oxygen interactions available with aspartic and glutamic acids (OA type)

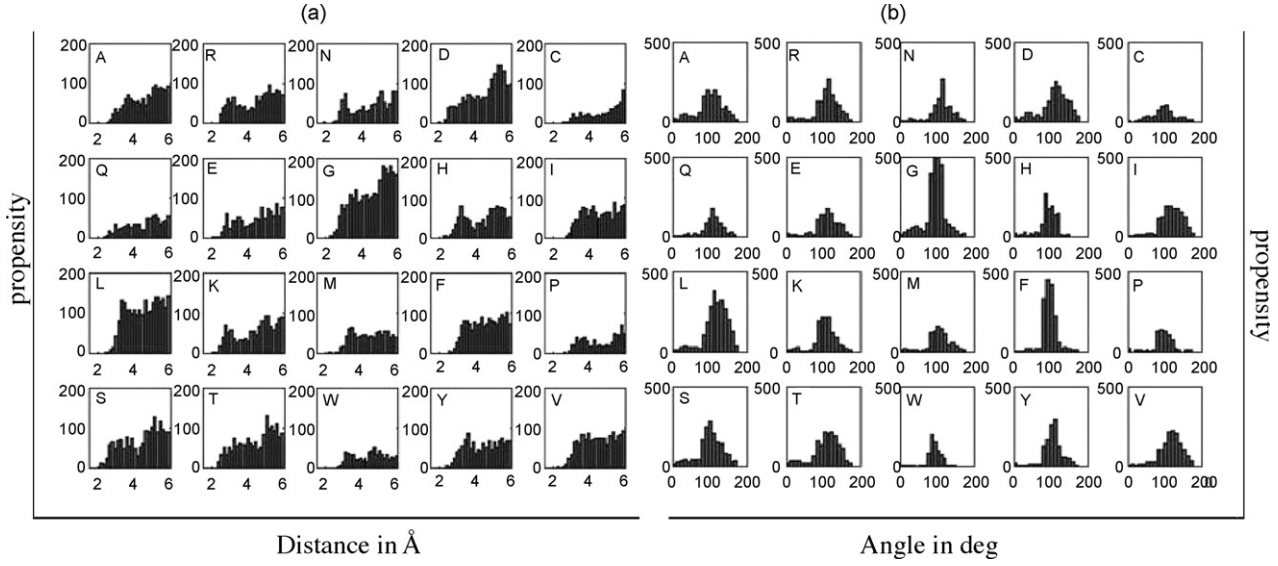

Fig. 2. Histograms that map the interactions between fluorine and the amino acids culled from protein–ligand complexes. The amino acids are labeled using their one-letter code. Panel "a" depicts the distances; panel "b" depicts the angles.

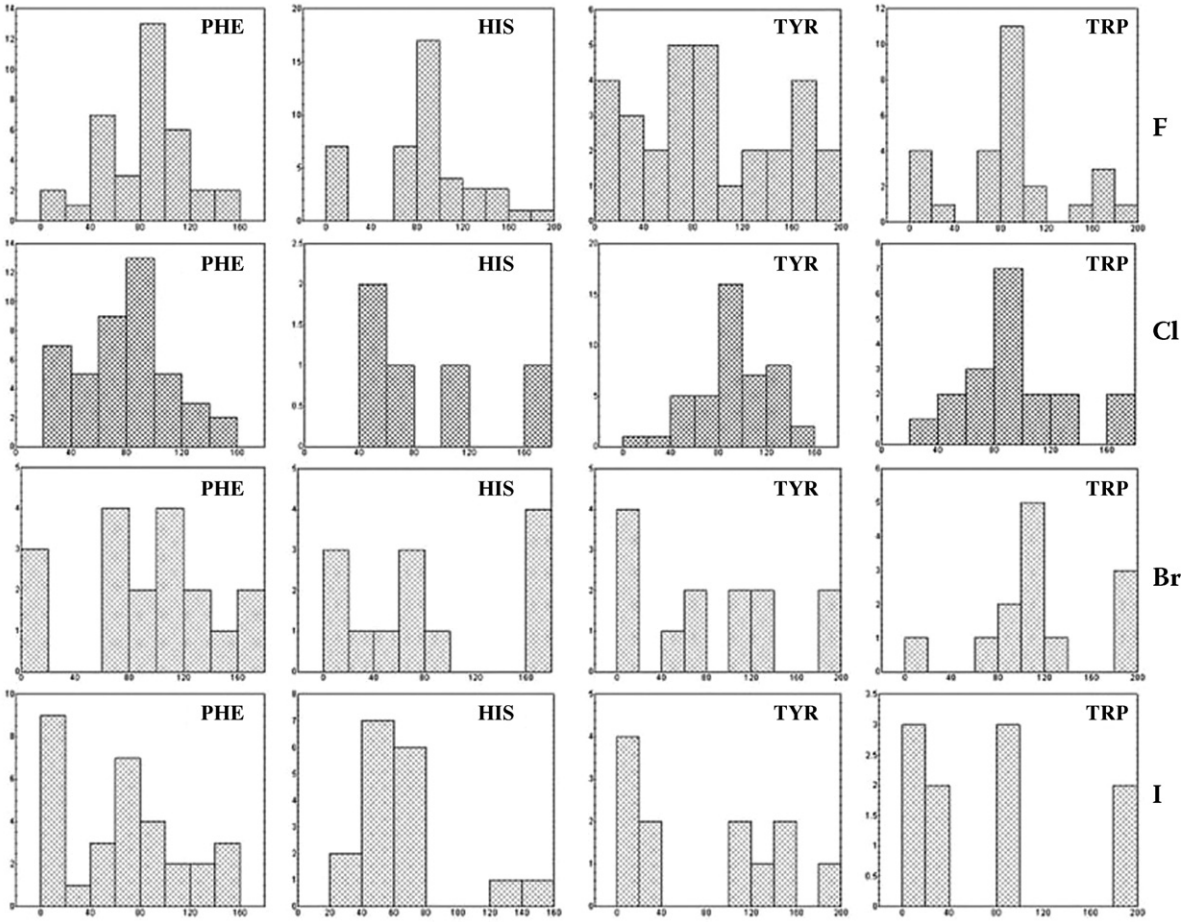

Fig. 3. Histogram representation of the interplanar angles between aryl-X and aromatic amino acids (marked in three letter code), where X = fluorine (F), chlorine (Cl), bromine (Br), or iodine (I).

but not serine or threonine (OL-type). Similarly, the distance distributions of aryl-F interactions with both OL and OA oxygens were bimodal with angles in the range of ∼100°, while the distance distributions of alkyl-F interactions with the two types of oxygens were unimodal with a peak angle of ∼120°. The bimodal distributions commonly found in the aryl-F interactions may be attributable to the presence of multiple fluorine groups such as geminal or trifluoromethyl groups. Halogen atoms, including those present in many toxic chemicals, exhibit a strong tendency to form halogen-bonded interactions in the presence of strong donor atoms (Fig. 4). Halogen-bonded interactions can also be found between the bromine atoms of both bromoform and 4-bromophenol interacting with glutamic acid in the binding site of the firefly luciferase enzyme [27] (Fig. 4B) and toluene monooxygenase [28] (Fig. 4C). The chlorine atoms of chloroacetone and dichloro propane form hydrogen bonds with trptophan and serine of hydronitrile lyase that has been implicated in cyanogenesis [29] (Fig. 4D). While the bromine atoms of dibromomethane interact with mainly hydrophobic residues of methane monooxygenase hydrolase [30] (Fig. 4E).

Water molecules are both multiple hydrogen bond donors and acceptors, and can thus act as bridges between interacting residues and/or ligands in proteins. However, many such interactions go undetected due to the practical difficulties in resolving the electron density of water molecules in crystallographic studies. An example of an important water mediated halogen ligand–protein interaction is found in the interaction of the halogenated ligand with LinB, a haloalkane dehalogenase found in many bacterial strains such as *Sphingomonas paucimobilis* UT26 [31] (Fig. 4F). Certain halogenated industrial pollutants such as 1,2-dichloroethane and 1,2-dichloropropane can bind to this enzyme and inhibit its activity. Structural studies on this protein–ligand complex by Oakley et al. [32] have shown that the inactivity of the enzyme in the presence of water mediated halogen interactions is a key factor in the biological activity of these pollutants. In our study we found that the interaction parameters of halogens with the oxygen of water were similar to those for other oxygen-containing amino acid residues. The mean distance of interaction between the various halogen atoms and oxygen atom of water varied from 2.6 to 3.6 Å, while the peak angle value was ∼110°. Analysis of the interactions of each of the halogen types with their hydrogen bonding amino acids and water molecules indicated that bromine and iodine, due to their weaker electronegativity, could be considered better leaving groups and hence should be more toxic. This suggestion is consistent with the finding that addition of bromine atoms to haloesters increased the toxicity of the compound [24].

Our propensity analysis shows that halogens do not favor interactions with sulfur-containing amino acids but, nevertheless, there are examples of sulfur-containing amino acids playing a key role as hydrogen bonding partners for halogens. One of the well-studied cases of sulfur–halogen interactions is with thyroxine in human serum albumin [33]. The iodine atoms of the ligand interact with sulfur-containing amino acids in the binding site. We found that most of the interactions with sulfur were with alkyl-substituted ligands, at a peak distance of 3.7 Å and a peak angle of ∼110°.

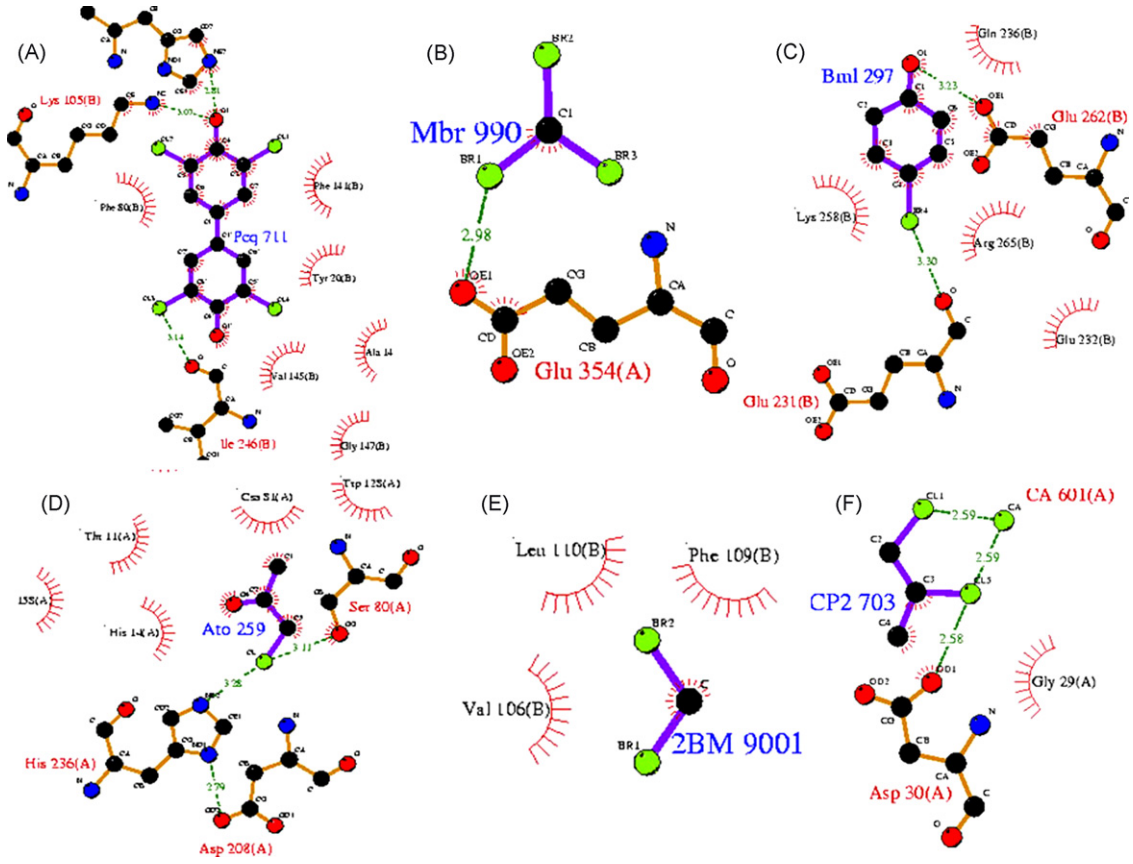

Fig. 4. Schematic LIGPLOT representation of the interaction between halogen atoms of ligands and amino acids in biological systems. Halogenated ligands are represented as ball-and-stick and colored-atom type (carbon = black, oxygen = red, nitrogen = blue, halogen = green); (A) polychlorinated biphenyl (PCB) is bound to the estrogen binding site of human estrogen sulfotransferase; (B) interaction of bromoform with firefly luciferase enzyme; (C) 4-bromophenol bound to toluene/*o*-xylene monooxygenase hydroxylase; (D) interaction of chloroacetone with hydroxynitrile lyase; (E) dibromomethane interaction with methane monooxygenase hydrolase; (F) interaction of dichloro propane with LINB, a haloalkane dehalogenase enzyme.

From our analysis we found several case studies in which increasing the number of halogens and or introducing halogens at key positions on an aryl ring had a strong impact on the activity of the ligand [34]. A recent study by Carosati et al. [35] described the nature and geometry of hydrogen bonds with different types of fluorine atoms in biological systems. Pesenti and Viani [36] described the effective use of fluorinated analogs in intercepting insect communication systems. The SAR for interactions of the pheromone receptor includes a polar functional group, a carbon–carbon double bond and the terminal alkyl group. All three elements of the pharmacophore have been subject to halogenation on a number of pheromones, such as (*Z*)-11-hexadecenyl acetate and (*E*)-11-tetradecenyl acetate, that mimic the natural pheromone of the insect. Results from these studies indicate that the mono-fluoro analogs of these compounds were able to mimic the natural pheromones. The other halogens in the series displayed poor inhibitory activities, suggesting that steric factors should be considered in the optimal design of the halogen analogs. Fluorination of the double bond was found to cause either enhancement or preservation of the natural pheromone activity; however, the effect of difluorination was highly position dependent. A difluromethylene group adjacent to the double bond may affect both the conformational flexibility and the double bond reactivity. The study also found that introduction of a single fluorine atom in the alkyl chain did not have a significant effect, while substitution of a bulky trifluoro group to the terminal alkyl group altered the binding affinity.

Chlorinated aromatics such as PCBs, chlorophenols, and the highly toxic 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) are commonly found in water. PCBs can exert reproductive toxicity by binding to the estrogen receptor [37] (Fig. 4A) or by inhibiting estrogen catabolism [38]. They are also known to influence transcriptional activation mediated by the thyroid hormone [39]. The chemical structure of highly chlorinated PCBs has been shown to play a key role in the slow metabolism of these compounds and, hence, their potential to disrupt major signaling pathways [40]. The nature, number and position of halogen atoms in PCBs make them a specific toxin to different species (e.g., mouse, rat and human). PCBs with 5–10 chlorine substituents activated the mouse and rat pregnane xenobiotic receptor (PXR), while the unsubstituted biphenyls were unable to activate either the rodent or human PXR (hPXR) implicating specific toxic effects of halogenation [41].

The distance distribution for fluorine interactions with all the amino acids shown in Fig. 2 was mostly bimodal with the first peak at ~3 Å and another peak at ~4.5 Å. A similar trend was observed among the other halogens. From these distributions it is evident that both short-range and long-range interactions are important. The closest distance between the halogen in the ligand and the amino acid was also determined. While no correlation was found between the closest distance of interaction and the size of the halogen, steric factors may play a role in accommodating the ligand in the binding site. A recent study by Chou et al. [42] described the SAR of benzothiophene anthranilamide class of compounds with Factor Xa (fX_a). The crystal structure of the ligand complexed with fX_a showed that the effects of substituents at various positions on the aryl ring of the ligand were highly influenced by the steric

properties. They concluded that substitution at the C-5 position with a halogen or methyl group was highly critical for fX_a potency, while substitution at the 4th position of the aniline ring with either the chloro- or bromo-groups significantly increased the binding affinity of the analogues.

The interplanar angle between the halogen-substituted aryl and the aromatic amino acids is yet another parameter that appears to influence the energetics of interactions. Many theoretical and experimental reports have emphasized the energetics of edge-to-face and stacking interactions in molecular recognition [43–46]. In the present study we found a population shift from edge-to-face to stacking type of interactions among aryl-X ligands in going down the periodic table from fluorine to iodine (Fig. 3). This trend coincides with the decrease in electronegativity from fluorine to iodine. Our evaluation shows a general shift in the population of molecules involved in edge-to-face interactions to a stacking interaction between aromatic rings. In evaluating this feature it is necessary to consider the role of the entire molecule in positioning the ligand in the binding site. Thus there could be cases of forced π stacks, or the most commonly found displaced π stacks of halogenated rings with aromatic rings of amino acids that could still be energetically favorable. An example of the use of such interactions in drug design is a recent study on Dopamine D_4 selective ligands [47]. Other features that could influence the binding affinity among aryl halogens and aromatic amino acids are the number of halogens on the ring and the position of halogenation, with reference to other groups on the ring. Thus in a study by Kim et al. [34] on the effect of fluorination of carbonic anhydrase CA-II inhibitor *N*-(4-sulfamylbenzoyl)benzylamine, a 10-fold change in affinity was observed that was highly dependent on the number of fluorines on the ring. This increase in affinity was attributed to a combination of dipole-induced dipole, dipole–quadrapole and quadrapole–quadrapole interactions between the fluorinated ring of the ligand and the two key residues phe131 and pro202. The interesting feature of this interaction is the use of intermediate-range interactions in the distance range of 4–6 Å between the aromatic amino acid and the halogen substituted to the aryl ring of the ligand in the design of high affinity inhibitors.

In summary, our study offers the first systematic and comprehensive analysis of the interactions of halogenated ligands with various amino acids, based on the July 2006 PDB available at the time of analysis. Various parameters such as distance, angle and interplanar angle derived from this study can be used to improve knowledge-based potentials for docking and scoring halogenated ligands. Many molecular mechanics force fields currently in use lack interaction parameters for halogens, even though the published literature provides ample evidence of the significance of halogens in pharmaceutical drug design. The ample structural information provided in the present study is crucial for deriving these parameters for halogen atoms for use in molecular mechanics force fields. By taking advantage of the special properties of halogens, selective and high affinity ligands can be computationally designed. We have also illustrated that halogens, though widely recognized for their beneficial effects in increasing the lipophilicity of drug molecules, can be hidden toxicophores. By understanding the geometric features of structure–toxicity interactions of the halogen atoms such as distance, angle and other structural parameters, it may be possible to design pharmaceuticals, pesticides and other chemicals that are environmentally benign.

Acknowledgments

The authors acknowledge access to the computational facilities at the Informatics Institute of UMDNJ and the academic computing services. Support for this work has been provided primarily by the US-EPA funded Environmental Bioinformatics and Computational Toxicology Center (ebCTC), under STAR Grant number GAD R 832721-010 to WJW. A similar analysis on an older PDB data set was presented as a poster at the 49th Annual Biophysical Society meeting by one of the authors (SK).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmgm.2008.04.001.

References

[1] M.B. van Niel, I. Collins, M.S. Beer, H.B. Broughton, S.K. Cheng, S.C. Goodacre, A. Heald, K.L. Locker, A.M. MacLeod, D. Morrison, C.R. Moyes, D. O'Connor, A. Pike, M. Rowley, M.G. Russell, B. Sohal, J.A. Stanton, S. Thomas, H. Verrier, A.P. Watt, J.L. Castro, Fluorination of 3-(3-(piperidin-1-yl)propyl)indoles and 3-(3-(piperazin-1-yl)propyl)indoles gives selective human 5-HT1D receptor ligands with improved pharmacokinetic profiles, J. Med. Chem. 42 (12) (1999) 2087–2104.
[2] M. Rowley, D.J. Hallett, S. Goodacre, C. Moyes, J. Crawforth, T.J. Sparey, S. Patel, R. Marwood, S. Patel, S. Thomas, L. Hitzel, D. O'Connor, N. Szeto, J.L. Castro, P.H. Hutson, A.M. MacLeod, 3-(4-Fluoropiperidin-3-yl)-2-phenylindoles as high affinity, selective, and orally bioavailable h5-HT (2A) receptor antagonists, J. Med. Chem. 44 (10) (2001) 1603–1614.
[3] P.K. Gillman, A review of serotonin toxicity data: implications for the mechanisms of antidepressant drug action, Biol. Psychiatry 59 (11) (2006) 1046–1051.
[4] Y.J. Wu, C.D. Davis, S. Dworetzky, W.C. Fitzpatrick, D. Harden, H. He, R.J. Knox, A.E. Newton, T. Philip, C. Polson, D.V. Sivarao, L.Q. Sun, S. Tertyshnikova, D. Weaver, S. Yeola, M. Zoeckler, M.W. Sinz, Fluorine substitution can block CYP3A4 metabolism-dependent inhibition: identification of (*S*)-*N*-[1-(4-fluoro-3- morpholin-4-ylphenyl)ethyl]-3-(4-fluorophenyl)acrylamide as an orally bioavailable KCNQ2 opener devoid of CYP3A4 metabolism-dependent inhibition, J. Med. Chem. 46 (18) (2003) 3778–3781.
[5] S.B. Rosenblum, T. Huynh, A. Afonso, H.R. Davis Jr., N. Yumibe, J.W. Clader, D.A. Burnett, Discovery of 1-(4-fluorophenyl)-(3*R*)-[3-(4-fluorophenyl)-(3*S*)-hydroxypropyl]-(4*S*)-(4-hydroxyphenyl)-2-azetidinone (SCH 58235): a designed, potent, orally active inhibitor of cholesterol absorption, J. Med. Chem. 41 (6) (1998) 973–980.
[6] R. Bartzatt, Two lead drug designs based on chloramphenicol as the parent structure, which express alkylation activity with potential for clinical applications, J. Pharm. Pharmacol. 55 (5) (2003) 653–660.
[7] S. Ali, I. Ahmad, A. Peters, G. Masters, S. Minchey, A. Janoff, E. Mayhew, Hydrolyzable hydrophobic taxanes: synthesis and anti-cancer activities, Anticancer Drugs 12 (2) (2001) 117–128.
[8] E.L. Thomas, P.M. Bozeman, M.M. Jefferson, C.C. King, Oxidation of bromide by the human leukocyte enzymes myeloperoxidase and eosinophil peroxidase, J. Biol. Chem. 270 (1995) 2906–2913.
[9] H. Ali, J. Rousseau, J.E. van Lier, 7 Alpha-methyl- and 11 beta-ethoxy-substitution of [125I]-16 alpha-iodoestradiol: effect on estrogen receptor-mediated target tissue uptake, J. Med. Chem. 36 (2) (1993) 264–271.
[10] K.R. Gee, E.A. Archer, L.A. Lapham, M.E. Leonard, Z.L. Zhou, J. Bingham, Z. Diwu, New ratiometric fluorescent calcium indicators with moderately attenuated binding affinities, Bioorg. Med. Chem. Lett. 10 (14) (2000) 1515–1518.
[11] Y. Tang, G. Ghirlanda, N. Vaidehi, J. Kua, D.T. Mainz, W.A. Goddard III, W.F. DeGrado, D.A. Tirrell, Stabilization of coiled-coil peptide domains by introduction of trifluoroleucine, Biochemistry 40 (9) (2001) 2790–2796.
[12] K.H. Lee, H.Y. Lee, M.M. Slutsky, J.T. Anderson, E.N. Marsh, Fluorous effect in proteins: de novo design and characterization of a four-alpha-helix bundle protein containing hexafluoroleucine, Biochemistry 43 (51) (2004) 16277–16284.
[13] J.H. Bae, P. Paramita Pal, L. Moroder, R. Huber, N. Budisa, Crystallographic evidence for isomeric chromophores in 3-fluorotyrosyl-green fluorescent protein, Chembiochem 5 (5) (2004) 720–722.
[14] T. Steiner, Hydrogen-bond distances to halide ions in organic and organometallic crystal structures: up-to-date database study, Acta Cryst. B54 (1998) 456–463.
[15] J.D. Dunitz, Organic fluorine: odd man out, Chembiochem 5 (5) (2004) 614–621.
[16] L. Shimoni, J.P. Glusker, The geometry of intermolecular interactions in some crystalline fluorine-containing organic compounds, Struct. Chem. 5 (1994) 383–397.
[17] P.K. Thallypally, A. Nangia, A Cambridge structural database analysis of the C–H···Cl interactions: C–H···Cl- and C–H···Cl–M often behave as hydrogen bonds but C–H···Cl–C is generally a van der Waals interactions, Cryst. Eng. Commun. 3 (2001) 114–119.
[18] P. Auffinger, F.A. Hays, E. Westhof, P.S. Ho, Halogen bonds in biological molecules, Proc. Natl. Acad. Sci. U.S.A. 101 (48) (2004) 16789–16794.
[19] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, The Protein Data Bank, Nucleic Acids Res. 28 (2000) 235–242.

[20] I.K. McDonald, J.M. Thornton, Satisfying hydrogen bonding potential in proteins, J. Mol. Biol. 238 (5) (1994) 777–793.
[21] J.A. Olsen, D.W. Banner, P. Seiler, B. Wagner, T. Tschopp, U. Obst-Sander, M. Kansy, K. Muller, F. Diederich, Fluorine interactions at the thrombin active site: protein backbone fragments H–C (alpha)–C═O comprise a favorable C–F environment and interactions of C–F with electrophiles, Chembiochem 5 (5) (2004) 666–675.
[22] A. Hoffmann-Roder, E. Schweizer, J. Egger, P. Seiler, U. Obst-Sander, B. Wagner, M. Kansy, D.W. Banner, F. Diederich, Mapping the fluorophilicity of a hydrophobic pocket: synthesis and biological evaluation of tricyclic thrombin inhibitors directing fluorinated alkyl groups into the p pocket, ChemMedChem 11 (2006) 1205–1215.
[23] J. Olsen, P. Seiler, B. Wagner, H. Fischer, T. Tschopp, U. Obst-Sander, D.W. Banner, M. Kansy, K. Muller, F. Diederich, A fluorine scan of the phenylamidinium needle of tricyclic thrombin inhibitors: effects of fluorine substitution on pK_a and binding affinity and evidence for intermolecular C–F···CN interactions, Org. Biomol. Chem. 2 (9) (2004) 1339–1352.
[24] A.D. DeWeese, T.W. Schultz, Structure–activity relationships for aquatic toxicity to Tetrahymena: halogen-substituted aliphatic esters, Environ. Toxicol. 16 (1) (2001) 54–60.
[25] K.S. Akers, G.S. Sinks, T.W. Schultz, Structure–toxicity relationships for selected halogenated aliphatic chemicals, Environ. Toxicol. Pharmacol. 7 (1999) 33–39.
[26] S. Shevtsov, E.V. Petrotchenko, L.C. Pedersen, M. Negishi, Crystallographic analysis of a hydroxylated polychlorinated biphenyl (OH-PCB) bound to the catalytic estrogen binding site of human estrogen sulfotransferase, Environ. Health Perspect. 111 (7) (2003) 884–888.
[27] N.P. Franks, A. Jenkins, E. Conti, W.R. Lieb, P. Brick, Structural basis for the inhibition of firefly luciferase by a general anesthetic, Biophys. J. 75 (5) (1998) 2205–2211.
[28] M.H. Sazinsky, J. Bard, A. Di Donato, S.J. Lippard, Crystal structure of the toluene/o-xylene monooxygenase hydroxylase from pseudomonas stutzeri OX1: insight into the substrate specificity, substrate channeling, and active site tuning of multicomponent monooxygenases, J. Biol. Chem. 279 (29) (2004) 30600–30610.
[29] H. Lauble, S. Forster, B. Miehlich, H. Wajant, F. Effenberger, Structure of hydroxynitrile lyase from Manihot esculenta in complex with substrates acetone and chloroacetone: implications for the mechanism of cyanogenesis, Acta Crystallogr. Sect. D 57 (2001) 194–219.
[30] D.A. Whittington, A.C. Rosenzweig, C.A. Frederick, S.J. Lippard, Xenon and halogenated alkanes track putative substrate binding cavities in the soluble methane monooxygenase hydroxylase, Biochemistry 40 (2001) 3476–3482.
[31] Y. Nagata, K. Miyauchi, J. Damborsky, K. Manova, A. Ansorgova, M. Takagi, Purification and characterization of a haloalkane dehalogenase of a new substrate class from a gamma-hexachlorocyclohexane-degrading bacterium, *Sphingomonas paucimobilis* UT26, Appl. Environ. Microbiol. 63 (9) (1997) 3707–3710.
[32] A.J. Oakley, Z. Prokop, M. Bohac, J. Kmunicek, T. Jedlicka, M. Monincova, I. Kuta-Smatanova, Y. Nagata, J. Damborsky, M.C. Wilce, Exploring the structure and activity of haloalkane dehalogenase from *Sphingomonas paucimobilis* UT26: evidence for product- and water-mediated inhibition, Biochemistry 41 (15) (2002) 4847–4855.
[33] I. Petitpas, C.E. Petersen, C.E. Ha, A.A. Bhattacharya, P.A. Zunszain, J. Ghuman, N.V. Bhagavan, S. Curry, Structural basis of albumin–thyroxine interactions and familial dysalbuminemic hyperthyroxinemia, Proc. Natl. Acad. Sci. U.S.A. 100 (11) (2003) 6440–6445.
[34] C.Y. Kim, J.S. Chang, J.B. Doyon, T.T. Baird Jr., C.A. Fierke, A. Jain, D.W. Christianson, Contribution of fluorine to protein–ligand affinity in the binding of fluoroaromatic inhibitors to carbonic anhydrase II, J. Am. Chem. Soc. 122 (2000) 12125–12134.
[35] E. Carosati, S. Sciabola, G. Cruciani, Hydrogen bonding interactions of covalently bonded fluorine atoms: from crystallographic data to a new angular function in the GRID force field, J. Med. Chem. 47 (21) (2004) 5114–5125.
[36] C. Pesenti, F. Viani, The influence of fluorinated molecules (semiochemicals and enzyme substrate analogues) on the insect communication system, Chembiochem 5 (5) (2004) 590–613.
[37] E.C. Bonefeld-Jorgensen, H.R. Andersen, T.H. Rasmussen, A.M. Vinggard, Effect of highly accumulated polychlorinated biphenyl congeners on estrogen and androgen receptor activity, Toxicology 158 (2001) 141–153.
[38] M.H. Kester, S. Bulduk, D. Tibboel, W. Meinl, H. Glatt, C.N. Falany, M.W. Coughtrie, A. Bergman, S.H. Safe, G.G. Kuiper, A.G. Schuur, A. Brouwer, T.J. Visser, Potent inhibition of estrogen sulfotransferase by hydroxylated PCB metabolites: a novel pathway explaining the estrogenic activity of PCBs, Endocrinology 141 (5) (2000) 1897–1900.
[39] T. Iwasaki, W. Miyazaki, A. Takeshita, Y. Kuroda, N. Koibuchi, Polychlorinated biphenyls suppress thyroid hormone-induced transactivation, Biochem. Biophys. Res. Commun. 299 (3) (2002) 384–388.
[40] J.T. Borlakoglu, J.P. Wilkins, Metabolism of di-, tri-, tetra-, penta- and hexachlorobiphenyls by hepatic microsomes isolated from control animals and animals treated with Aroclor 1254, a commercial mixture of polychlorinated biphenyls (PCBs), Comp. Biochem. Physiol. C 105 (1) (1993) 95–106.
[41] M.M. Tabb, V. Kholodovych, F. Grun, C. Zhou, W.J. Welsh, B. Blumberg, Highly chlorinated PCBs inhibit the human xenobiotic response mediated by the steroid and xenobiotic receptor (SXR), Environ. Health Perspect. 112 (2) (2004) 163–169.
[42] Y.L. Chou, D.D. Davey, K.A. Eagen, B.D. Griedel, R. Karanjawala, G.B. Phillips, K.L. Sacchi, K.J. Shaw, S.C. Wu, D. Lentz, A.M. Liang, L. Trinh, M.M. Morrissey, M.J. Kochanny, Structure–activity relationships of substituted benzothiophene-anthranilamide factor Xa inhibitors, Bioorg. Med. Chem. Lett. 13 (3) (2003) 507–511.
[43] A.C. Hunter, J.K.M. Sanders, The nature of pi–pi interactions, J. Am. Chem. Soc. 112 (1990) 5525–5534.
[44] S.K. Burley, G.A. Petsko, Aromatic–aromatic interaction: a mechanism of protein structure stabilization, Science 229 (4708) (1985) 23–28.
[45] E.A. Meyer, R.K. Castellano, F. Diederich, Interactions with aromatic rings in chemical and biological recognition, Angew. Chem. Int. Ed. Engl. 42 (11) (2003) 1210–1250.
[46] H. Adams, J.L. Blanco, G. Chessari, C.A. Hunter, C.M. Low, J.M. Sanderson, J.G. Vinter, Quantitative determination of intermolecular interactions with fluorinated aromatic rings, Chemistry 7 (16) (2001) 3494–3503.
[47] S. Kortagere, P. Gmeiner, H. Weinstein, J.A. Schetz, Certain 1,4-disubstituted aromatic piperidines and piperazines with extreme selectivity for the dopamine D4 receptor interact with a common receptor microdomain, Mol. Pharmacol. 66 (6) (2004) 1491–1499.