

# **Halogen bonding at the ATP binding site of protein kinases: Preferred geometry and topology of ligand binding.**

Jarosław Poznański<sup>\*1</sup>, David Shugar<sup>\*1,2</sup>

<sup>[1]</sup> Institute of Biochemistry and Biophysics PAS, Pawińskiego 5a, 02-106 Warszawa, Poland

<sup>[2]</sup> Division of Biophysics, Institute of Experimental Physics, University of Warsaw, 93 Żwirki i Wigury St., 02-089 Warszawa, Poland

\*Corresponding authors: David Shugar, Jarosław Poznański, Institute of Biochemistry and Biophysics PAS, Pawińskiego 5a, 02-106 Warszawa POLAND, Tel +48 22 592 3511, Fax: +48 22 592 2190, E-mail: [shugar@ibb.waw.pl](mailto:shugar@ibb.waw.pl), [jarek@ibb.waw.pl](mailto:jarek@ibb.waw.pl)

## ABSTRACT

Halogenated ligands have been widely developed as potent, and frequently selective, inhibitors of protein kinases (PK). Herein, all structures of protein kinases complexed with a halogenated ligand, identified in the PDB, were analyzed in the context of eventual contribution of halogen bonding to protein-ligand interactions. Global inspection shows that two carbonyl groups of residues located in the hinge region are the most abundant halogen bond acceptors. In contrast to solution data, well-defined water molecules, located at sites conserved across most PK structures, are also involved in halogen bonding. Analysis of cumulative distributions of halogen-acceptor distances shows that structures displaying short contacts involving a halogen atom are overpopulated, contributing together to clearly defined maxima of 2.82, 2.91 and 2.94 Å for chlorine, bromine and iodine, respectively. The angular preference of a halogen bond favors ideal topology (180°, 120°) for iodine. For bromine the distribution is much more dispersed, and no such preference was found for chlorine.

## INTRODUCTION

Specific interactions between the halogen atoms (Cl, Br, I) of halogenated ligands and the electron pairs of oxygen/nitrogen/sulfur have been identified in many crystal structures of supramolecular ensembles [1-3] as well as in complexes between biomolecules and halogenated ligands [4-8]. In both cases this is based in part on the observation that the distance between a halogen atom and its electron-donating partner is significantly lower than the sum of their vdW radii (Figure 1). Fluorine, because of its high electronegativity, is a very poor halogen bond acceptor, but is a good acceptor from standard hydrogen bond donors [9].

Figure 1

Bearing in mind that numerous natural drugs, and an increasing number of synthetic drug candidates, are halogenated, understanding the nature and thermodynamics of halogen bonding should contribute to rational drug design. Halogenated compounds comprise 50% of current screening libraries [10], and almost 20% of low-mass protein ligands listed in the PDB are halogenated. Current widespread interest in elucidating the role of halogenated ligands in biological systems has been widely reviewed, amongst others, by Voth & Ho [7] and Parisini et al. [5], demonstrating that halogen bonding interactions are responsible for ligand recognition in many protein structures deposited in the PDB. Statistical analysis showed that such halogen bonds exhibit a geometry similar to that of hydrogen bonds [4], including a directional preference. However, despite the clear evidence of short halogen-bonding contacts in many protein-ligand structures, their energetic contribution is still under active debate [2, 6, 11-19]. Estimated contributions of halogen bonding to intra- and intermolecular interactions in aqueous medium vary from 0.2 [2] up to 5-8 kcal/mol [11]. The apparently largest values for X-bond contributions in solution were estimated, with the aid of IR spectroscopy, for  $\text{CF}_3\text{-X}\dots\text{N}(\text{CH}_3)_3$  in liquid noble gases, as 2.1, 4.4 and 6.8 kcal/mol for X-bonds involving Cl, Br and I, respectively [13]. However, the fundamental problem concerning the relation between hydrophobic and electrostatic interactions, and the enthalpy of hydrogen vs. halogen bond formation, remains to be resolved. In most experimental studies, in which the free energy of interaction of a ligand is usually compared to that of its halogenated counterpart, direct halogen bonding is assumed to be a major factor contributing to the observed differences in binding. Other interactions (e.g. hydrogen bonding, electrostatic interactions, and hydrophobic contribution) are frequently neglected.

One of the first reported inhibitors of the ubiquitous protein kinase CK2 [20], relatively selective when tested against a panel of more than 60 other kinases [21, 22], is 4,5,6,7-tetrabromobenzo-1H-triazole (TBBt, see Figure 2). Other structurally closely related brominated analogues, TBBz (4,5,6,7-tetrabromo-1H-imidazole) [23], K59 (4,5,6,7-tetrabromo-1H-indazole) and K6X (3,4,5,6,7-pentabromo-1H-indazole) [24] have been synthesized, and displayed even higher activity against CK2 $\alpha$  [22, 24, 25]. Although TBBt, TBBz and K59 have very similar structures (see Figure 2), differing only by replacement of a single nitrogen in the triazole ring of TBBt by a C(2) or C(3) carbon in TBBz and K59, respectively, they differ significantly in physico-chemical properties, especially their pKa values (4.95, 8.05 and 10.6 for TBBt, TBBz and K59, respectively). Furthermore, all three are sparingly soluble in neutral aqueous medium.

Figure 2

The foregoing differences in the crystal structures clearly indicate that binding of a halogenated ligand is driven by a balance of electrostatic, hydrophobic, hydrogen-bonding and halogen-bonding interactions, with the electrostatic component being predominant. This is supported by analysis of the structure-activity relations for a series of 4,6,7-tribromobenzotriazole derivatives [26], in which the Br at C(5) of TBBt is replaced by various groups differing in their physicochemical properties, and for another series of nine bromobenzo-1H-triazoles representing all possible patterns of halogenation of the benzene ring [27]. Overall, the hydrophobicity of the monoanionic form of the ligand appeared to be the principal factor governing its inhibitory activity against CK2 $\alpha$  [26, 28]. Furthermore, the moderate inhibitory activity exhibited by 4,5,6,7-tetramethylbenzotriazole [20], which, in contrast to TBBt, is in the neutral form at physiological pH [29], again points to a balance of electrostatic and hydrophobic interactions as important factors contributing to CK2 $\alpha$  inhibition.

Herein we present a detailed analysis of the geometry and topology of halogen bonds identified in all complexes of protein kinases with halogenated ligands accessible in the Protein Data Bank. A statistical approach was applied to estimate, independently for Cl, Br and I as halogen bond donors, the free energy of halogen bond formation in protein-ligand systems, with some emphasis on protein kinase CK2.

## Material and methods

### *Structural data*

The PDB [30] was searched to identify all entries of protein kinases (enzyme classifications (EC) 2.7.10: tyrosine kinase; 2.7.11: serine/threonine kinase; 2.7.12: dual-specificity kinase; 2.7.13: histidine kinase). Those containing ligands with at least one X-carbon bond (X = Cl, Br, I) were then analyzed.

### ***Structural analysis***

Analyses were performed with the aid of the Yasara Model package [31]. For each halogen bond donor (X), potential acceptors (Acc) were identified in a sphere of 4 Å radius. All protein oxygen atoms (i.e. backbone carbonyls, side-chain carbonyls of Asn and Gln, carboxyls of Asp and Glu, and hydroxyls of Ser and Thr), nitrogen atoms (backbone, and those of Arg, His and Lys), and side-chain sulphurs of Met and Cys, were assumed to be potential halogen bond acceptors. Following the concept of halogen bonding to  $\pi$ -electron systems [32], the centers of the aromatic rings of His, Phe, Tyr and Trp were also analyzed. Water molecules were also taken into account as potential halogen bond acceptors, albeit both experimental and theoretical data suggest that a water molecule is not an efficient halogen-bond acceptor. Binding of halide anions (Cl<sup>-</sup>, Br<sup>-</sup> and I<sup>-</sup>), for which electrostatic interactions predominate, was excluded from the analysis.

Structures of all protein kinases were aligned, using the modified MUSTANG algorithm [33] implemented in Yasara Structure. Multiple protein molecules in the crystal cell, as well as objects displaying partially occupied forms (i.e. side-chain rotamers or ligand locations) were analyzed separately.

### ***Statistical analysis***

Calculated cumulative distributions were based on all identified X...Acc distances. The superposition of one, two or three cumulative distribution functions (CDF) of the normal distribution, each with a mean  $\mu_i$  and a standard deviation  $\sigma_i$ , was fitted to experimental data with the aid of the Marquardt algorithm [34] implemented in the GnuPlot program [35] according to the formula.

$$CDF(d) = 0.5 \cdot \sum_i (1 + erf[(d - \mu_i) / \sqrt{2} / \sigma_i^2])$$

The CDF analysis was restricted to X...Acc distances (d) smaller than the sum of the vdW radii of the X and Acc atoms (3.3, 3.4 and 3.5 Å for X = Cl, Br and I, respectively).

## **RESULTS**

### *PDB screening.*

A total of 334 entries were identified, among them 249, 73 and 33 with chlorinated, brominated and iodinated ligands, respectively. These entries were analyzed in terms of short contacts (4 Å threshold) between a halogen atom and a potential halogen bond acceptor (Acc: oxygen, nitrogen, sulfur and the center of the  $\pi$ -electron system of His, Phe, Tyr and Trp).

In 69 structures no contacts were observed. In the remaining ones, 1131 short X...Acc contacts were identified (see Supplementary Table S1 for the list of all protein kinases identified). This includes 816 C-Cl...Acc, 260 C-Br...Acc and 55 C-I...Acc interactions. For each the X...Acc distance, and C-X...Acc and X...Acc-C angles (the latter with exception of a water molecule and a  $\pi$ -electron system), were determined to define the spatial preferences for formation of a halogen bond. Since the vdW radii of Cl, Br and I differ noticeably (1.75, 1.85 and 1.98 Å, respectively), the three sets of contacts were analyzed separately. Sample sizes (816, 260 and 55) precluded a standard approach of direct analysis of a probability distribution with reasonable bins' size [4]. Hence an alternative approach, involving analysis of a cumulative distribution, was applied.

### *Initial data verification*

All structures displaying abnormally short X...Acc distances were carefully inspected to trace unreported dehalogenation or other ambiguities. Thus, extremely short contacts between a halogen atom and a water molecule were identified in 3KXN (2.22 Å distance of I17<sup>K8X</sup>...HOH<sup>480</sup>), in 3PVG (2.59 Å for Br19<sup>K68</sup>...HOH<sup>439</sup>), in 3AD5 (2.16 Å for Cl49<sup>5PB</sup>...HOH<sup>154</sup>), in 2R3F (2.53 Å for Cl25<sup>SC8</sup>...HOH<sup>1205</sup>) and in 3LCD (2.66 Å for Cl17<sup>BDY</sup>...HOH<sup>14</sup>). All of these were considered too short, and removed from further analysis.

### *Analysis of X...Acc distance distribution*

Cumulative distributions for X...Acc distances, obtained independently for three data sets (Cl...Acc, Br...Acc and I...Acc), are presented in Figure 3.

Figure 3

The visible shoulders (see gray ellipses located close to 3Å), observed for all cumulative distributions, indicate that short contacts are overrepresented. This is most evident for the cumulative distribution of Br...Acc distances, in which the deviations in the short-distance region (i.e. below 3Å) explicitly define a remaining contribution from halogen bonding. Consequently, following the concept of halogen bonding, the CDF of two Gaussian distributions was fitted to the data (see dotted lines in Figure 3 A-C). At this stage, the CDF

function of superposition of two Gaussian distributions was found to reconstruct satisfactorily the cumulative distributions for X...Acc distances shorter than 3.30 and 3.40 Å for Cl and Br, respectively. These two values agree with the sums of oxygen and halogen vdW radii, 3.27 and 3.37 Å for Cl...O and Br...O, respectively. For longer distances, the distance distribution depends rather on the spatial preference of binding sites than on the energetics of direct interaction between a halogen atom and a potential acceptor. The short-distance distributions fitted for Cl...Acc ( $\mu = 2.82$  Å,  $\sigma = 0.08$  Å, contribution 2%) and for Br...Acc ( $\mu = 2.91$  Å,  $\sigma = 0.02$  Å, 3%) are rough estimates of a contribution of halogen bonding in protein-ligand systems.

As regards I...Acc distances, the small sample size (54 identified contacts, 44 shorter than a vdW cutoff of 3.5 Å) makes even a CDF barely qualitative. However, the fit of the CDF function corresponding to superposition of two normal distributions displays visible deviations below 3 Å. These deviations can be compensated for by a contribution from the third Gaussian distribution centered close to 3 Å (see dotted line in Figure 3C). The explicit X...Acc distributions can be directly calculated as formal derivatives of the fitted CDF. Inspection of Figure 3D-F shows that the estimated distance distributions perfectly reconstruct, up to vdW thresholds (3.3, 3.4 and 3.6 Å), histograms obtained with 0.1 Å bins (gray bars denote distances shorter than assumed thresholds). This again confirms that the proposed method of analysis of a cumulative distribution is at least as informative as the commonly used direct analysis of distance distribution

In all cases the location of a short-distance local maximum follows the order of the vdW radii of halogen atoms ( $\mu = 2.82$ , 2.91 and 2.94 Å for Cl, Br and I). The reverse trend observed in distribution widths ( $\sigma = 0.08$ , 0.02 and 0.01 Å, respectively) confirms that the chlorine atom is the weakest halogen bond donor.

The estimated distributions of X...Acc distances can be converted to a thermodynamic semi-empirical potential of mean force (PMF), which can then be further used to verify the general force field parameterizations for halogen bonding [36-39]. These potentials, when determined separately for Cl...Acc, Br...Acc and I...Acc, show that the maximal stabilization of an X...Acc halogen bond is in the range of 0.5 kcal/mol for Cl to 1.5 kcal/mol for I. These values are approximately ten-fold lower than those reported by Voth et al [6], but agree with *ab initio* derived results for various X...Acc systems [40].

### ***Preferred geometry of a halogen bond.***

A total of 72 potential halogen bonds, unequivocally identified in 51 PDB structures of protein kinases, were inspected for a possible angular preference of halogen bonding (Figure 4).

Figure 4

In principle, an almost linear conformation of the C-X...Acc fragment is strongly favored (i.e.  $\theta_X \approx 180^\circ$ ), but two distinguishable maxima (visible as shoulders in the cumulative distribution) are also evident (see Figure 5A). One of them, centered close to  $\theta_X \approx 123^\circ$ , is characteristic for interaction of a halogen atom with a well-defined molecule. Thus, an isolated water molecule seems to be a strong halogen bond acceptor in numerous crystal structures of protein – ligand complexes, as represented by two ligand orientations in the complex of K37 with CK2 $\alpha$ , for which a water molecule competes with a carbonyl oxygen for a bromine donor (see Supplementary Figure S1). The third minimum, located at  $\theta_X \approx 160^\circ$ , also noted by Auffinger et al. [4], is an alternative to a linear halogen bond.

More detailed analysis shows that the angular preference strongly depends on the type of halogen atom. Thus, the linear form of the C-X...Acc configuration exists for almost 60% of short contacts involving Iodine, but only approximately 30% for Bromine, and no preference for a linear X-bond is observed for Chlorine. This is additional evidence that Iodine is the strongest halogen bond donor, and Chlorine the weakest.

The majority of X...Acc-C angles ( $\theta_{Acc}$ ) agrees with the expected location of a maximal electron density on a carbonyl. In three plots, clusters located at X...Acc-C ( $\theta_{Acc}$ ) =  $125^\circ$  and C-X...Acc ( $\theta_X$ ) =  $175^\circ$  represent those with ideal X-bond geometry. It is clearly visible that the strongest preference for optimal geometry of a halogen bond is that for Iodine (Figure 4D). The X-bond structures are more dispersed for Bromine (Figure 4C) and much more so for Chlorine (Figure 4B). Note that for a water molecule,  $\theta_{Acc}$  is undefined.

#### ***Preferred topology of halogen bond network in the ATP binding site of protein kinases.***

The largest number of short X...Acc contacts, with at least 3 halogen atoms engaged in X-bonds, are visible for TBBt complexed with CDK2 (PDB ID:1P5E [41]) and for K37 (PDB ID: 1ZOG [42]), TBBz (PDB ID: 2OXY [23]) and K6X (PDB ID: 3KXG [24]) complexes with CK2 $\alpha$  (see Supplementary Figure S1). In all these structures, and frequently in others, the backbone carbonyl oxygens of residues located in the hinge region (i.e. Glu114 and Val 116 in CK2 $\alpha$ , and Glu81 and Leu83 in CDK2), are the most favored acceptors (see Supplementary Figure S2 for all complexes with identified halogen bonding to the hinge region). Two additional halogen bonds may be created with proximal water molecules, one of which is well defined at locations conserved among crystal structures of protein kinases [23]. However, one of the water molecules can be replaced by the backbone carbonyl oxygen of CDK2 Ile10 (Val45 in CK2 $\alpha$ ), which precedes the phosphate anchor loop (p-loop, residues <sup>46</sup>GRGKYS<sup>51</sup> in



CK2 $\alpha$ ), and the other by a side-chain carboxyl of Asp175 (Asp145 in CDK2), which is located in the Mg<sup>2+</sup>-binding loop (<sup>175</sup>DWG<sup>177</sup> motif in CK2 $\alpha$ ).

In two of these crystal structures (PDB IDs: 1P5E, 2OXY), the patterns of halogen bonds remain identical for both protein molecules located in the crystal asymmetric unit. In two other structures (PDB IDs: 1ZOG, 3XKN) two alternative, partially occupied, locations of the ligand in the binding pocket are visible. However, in all these structures a pattern of close contacts of ligand halogen atoms with backbone carbonyls of residues located in the N-terminal part of the hinge region is preserved for three different orientations of the ligands. In the next two structures (PDB IDs: 3KXG, 1J91), the proteins display partial occupancy of two conformers, in one of which a halogen bond with the ligand is observed.

It should be emphasized that, despite differences in orientation of the ligand in the binding pocket, it is the pattern of halogen bonds that is preserved rather than the location of a ligand (Figure 5).

Figure 5 (STEREO).

Of 51 identified structures, 23 represent the most populated pattern of interactions (Figure 5, see also Supplementary Figure S2 and Supplementary Table S1), in which halogenated ligands make direct halogen bond contacts with the side-chain carbonyls of the N-terminal part of the kinase hinge/ $\alpha$ D region. In most of these structures, the carbonyl groups of residues located in the N-terminal part of the hinge region are halogen bond acceptors (PDB IDs: 1P5E and 3MY5 of CDK2; 3MY1 of CKD9; 1ZOE, 1ZOG, 1ZOH, 2OXD, 2OXX, 2OXY, 3KXG, 3KXM, 3KXN, 3PVG and 3OFM of CK2 $\alpha$ ; and 2VAG of CLK1). Halogen bonds may also be formed with other residues located in the hinge region (PDB IDs: 1CKP and 1H1R of CDK2; 1PMN and 1PMQ of JNK3; and 2ZOQ of ERK1), or even with a proximal water molecule (PDB ID: 1WCC of CDK2).

An alternative halogen bonding mode occurs to the p-loop region (Arg43-Phe54 in CK2 $\alpha$ ) in 1J91, 1ZLT, 2GU8, 2UW8, 2X6D 3NGA 3SDO, frequently accompanying hydrogen bonding between a ligand and backbone carbonyls of residues located in the hinge region (see Supplementary Figure S3). Other potential acceptors for halogen bonds are residues located in the DWG motif of the Mg<sup>2+</sup>-binding loop (Asp175-Gly177 in CK2 $\alpha$ ). Thus, the side-chain carboxyl of Asp (PDB IDs: 1ZOH, 3KXG, 1Q4L, 1FVT, 1UV5, 3ZZ2), a backbone carbonyl of Gly (PDB IDs: 2XP2, 2YFX), a side-chain hydroxyl of the Ser preceding the DWG motif (PDB IDs: 2Y6O, 3PJ1), or the  $\pi$ -electron system of Phe (PDB ID: 3SOA; in which F corresponds to W in the DWG motif of CK2 $\alpha$ ), also participate in halogen bonding (see Supplementary Figure S4).

## DISCUSSION

It should be noted that an unexpectedly large number of extremely short contacts ( $< 3\text{\AA}$ ) between a water molecule and a halogen atom were identified (PDB IDs: 1P5E, 1ZOE, 1ZOG, 2OXD, 2OXY, 3PVG, 1WCC, 2R3F, 2R3P, 2C47, 1S9J). This substantially disagrees with solution data [2, 43], and also with results of QM calculations [40], which predict that a water molecule is not a halogen bond acceptor. It could thus be concluded that the weak accepting properties of a water molecule result from the large entropic contribution upon formation of an eventual halogen bond. This unfavored term is diminished for a water molecule tightly bound to the protein, and thus a halogen bond could be formed in a ligand-water-protein ternary system. In this context, halogen bonding to an isolated water molecule should be considered in ligand-protein interactions as an alternative to water-bridged hydrogen bonding.

In two structures, 1P5E [41] and 2OXY [23], ligand locations in two complexes made with different protein molecules in the asymmetric unit remain identical. By contrast, two alternative locations of K37 complexed with CK2 $\alpha$  differ in orientation by  $60^\circ$ , but the pattern of two halogen bonds to Glu114 and Val116 is preserved (see Supplementary Figure S1). In the complex of K6X with CK2 $\alpha$ , two alternative rotamers of the Asp173 side-chain are partially populated, but only one makes a halogen bond with the ligand. Such variations in ligand orientation are observed for numerous ligands, clearly denoting only a minor contribution of halogen bonding to stabilization of the ligand-protein complex, additionally supported by differences in the pattern of halogen bonding within partially populated conformations, or in different molecules from the same crystal unit.

Inspection of crystal structures of TBBt, TBBz and K64 complexed with CK2 $\alpha$  [28] reveals significant differences in their location in the ATP binding site, as well as in the topology of the intermolecular halogen bonds. Thus, location of TBBz is constrained by non-polar contacts, and two halogen bonds between the backbone carbonyls of E114 and V116 with the vicinal Br5 and Br6 atoms (3.16 and 2.95  $\text{\AA}$ , respectively), whereas K64, which is rotated in the complex by  $\sim 60^\circ$ , retains a halogen bond with V116 (2.85  $\text{\AA}$ ) and a long-range halogen-bond-like contact with E114 ( $> 3.5\text{\AA}$ ), but also forms an additional halogen bond with the D175 side-chain in one of its partially occupied conformations (3.09  $\text{\AA}$ ). In striking contrast, location of TBBt is stabilized, apart from non-polar contacts, mainly by electrostatic interactions between the triazole ring and the side-chains of K68 and D175, and an orthogonal halogen bond, see [6] for a definition, between the N $\epsilon$  of R47 and a peripheral Br4/Br7 of TBBt (2.99  $\text{\AA}$ ), inadvertently overlooked by the authors [44]. All three ligands also make numerous

contacts with well-defined surrounding water molecules, the oxygen atoms of some of which are closer than the nominal limit for an O...Br halogen bond (3.3 Å) (see Figure 7)

There are also some ligands (e.g. CX-4945, the most promising CK2 $\alpha$  inhibitor, now in clinical trials) which are bound in a halogen bond-promoting mode, but distances between the halogen atom and proximal halogen bond acceptors are too long. Structural analysis demonstrates that in the complex of CX-4945 with CK2 $\alpha$ , the small hydrophobic cavity is large enough to host not only chlorine [45], but also the larger bromine, or even iodine, attached to the meta-position of the benzene ring of the parent CX-4945 inhibitor. Consequently, replacement of Cl by either Br or I increases the hydrophobic contribution to the free energy of binding, which may additionally be enforced by creation of a halogen bond with the carbonyl of Gly46. A preliminary *in silico* study, with the aid of a docking procedure, showed that replacement of Cl by Br increases the free energy of binding by 0.25 kcal/mol, and by 0.35 kcal/mol when Cl is replaced by I.

It should also be noted that halogenation of a ligand may change its pKa, and/or even the tautomeric/protomeric equilibrium experienced by proximal dissociable groups. This may result in modulation of hydrogen bonding interactions, which accompanies eventual halogen bonding. Thus, the reported activity of a series of halogenated PDE-5 inhibitors correlates with the calculated energies of halogen bonding with a tyrosine hydroxyl group [46], albeit the authors neglected, the significant contribution from two hydrogen bonds created between the ligand and a Gln817 side-chain, for which the estimated energies, based on crystal structures (PDB IDs: 3SHY, 3SHZ and 3SIE, are 25.3, 32.0 and 37.9 kJ/mol for fluoro-, chloro- and bromo- derivatives, respectively, visibly correlating with the reported IC<sub>50</sub> values ( $R^2=0.98$ )

## Figures

**Figure 1.** Schematic representation of a halogen bond

**Figure 2.** Effective multihalogenated inhibitors of CK2 $\alpha$ .

**Figure 3.** CDF analysis based on the assumption that the observed X...Acc distribution (represented as cumulative distribution) of a normal distribution (A, B, C, solid lines) or a superposition of two or three normal distributions (A, B, C dotted lines). Gray ellipses denote the significant deviations from a normal distribution. In bottom row a comparison of CDF-derived distributions with frequency histograms calculated with 0.1 Å bins. All X...Acc pairs separated by more than the sum of their vdW radii were excluded from the analysis.

**Figure 4.** Cumulative distributions of C-X...Acc angles for 72 identified halogen bonds obtained separately for different halogen bond donors (A), and scatterplots for C-X...Acc vs. X...Acc-C angles demonstrating geometrical preferences for halogen bonds involving Cl (B), Br (C) and I (D), respectively. Gray ellipses denote the most favored geometry of a C-X...O=C halogen bond.

**Figure 5.** Stereo representation of halogen bonds identified in the ATP-binding pocket of protein kinases. Gray neons represent individual halogen bonds, and green ones trace the protein backbone. Water molecule (Wat), which location is conserved among the crystal structures of various protein kinases, is engaged in halogen bonding similarly to backbone carbonyls from the hinge region.

## REFERENCES

- [1] P. Metrangolo, F. Meyer, T. Pilati, G. Resnati, G. Terraneo, Halogen bonding in supramolecular chemistry, *Angewandte Chemie-International Edition*, 47 (2008) 6114-6127.
- [2] M.G. Sarwar, B. Dragisic, L.J. Salsberg, C. Gouliaras, M.S. Taylor, Thermodynamics of Halogen Bonding in Solution: Substituent, Structural, and Solvent Effects, *Journal of the American Chemical Society*, 132 (2010) 1646-1653.
- [3] K. Rissanen, Halogen bonded supramolecular complexes and networks, *Cryst. Eng. Comm.*, 10 (2008) 1107–1113.
- [4] P. Auffinger, F.A. Hays, E. Westhof, P.S. Ho, Halogen bonds in biological molecules, *Proceedings of the National Academy of Sciences of the United States of America*, 101 (2004) 16789-16794.
- [5] E. Parisini, P. Metrangolo, T. Pilati, G. Resnati, G. Terraneo, Halogen bonding in halocarbon-protein complexes: a structural survey, *Chemical Society Reviews*, 40 (2011) 2267-2278.
- [6] A.R. Voth, F.A. Hays, P.S. Ho, Directing macromolecular conformation through halogen bonds, *Proceedings of the National Academy of Sciences of the United States of America*, 104 (2007) 6188-6193.
- [7] A.R. Voth, P.S. Ho, The role of halogen bonding in inhibitor recognition and binding by protein kinases, *Current Topics in Medicinal Chemistry*, 7 (2007) 1336-1348.
- [8] M.A.A. Ibrahim, Molecular Mechanical Study of Halogen Bonding in Drug Discovery, *Journal of computational chemistry*, 32 (2011) 2564-2574.
- [9] P. Politzer, J.S. Murray, M.C. Concha, Halogen bonding and the design of new materials: organic bromides, chlorides and perhaps even fluorides as donors, *Journal of Molecular Modeling*, 13 (2007) 643-650.
- [10] [http://iccb.med.harvard.edu/screening/compound\\_libraries/index.htm](http://iccb.med.harvard.edu/screening/compound_libraries/index.htm).
- [11] M. Carter, P.S. Ho, Assaying the Energies of Biological Halogen Bonds, *Crystal Growth & Design*, 11 (2011) 5087-5095.
- [12] W.S. Zou, J. Han, W.J. Jin, Concentration-Dependent Br center dot center dot center dot O Halogen Bonding between Carbon Tetrabromide and Oxygen-Containing Organic Solvents, *Journal of Physical Chemistry A*, 113 (2009) 10125-10132.
- [13] D. Hauchecorne, B.J. van der Veken, A. Moiana, W.A. Herrebout, The C-Cl ...N halogen bond, the weaker relative of the C-I and C-Br ...N halogen bonds, finally characterized in solution, *Chemical Physics*, 374 (2010) 30-36.
- [14] D.A. Kraut, M.J. Churchill, P.E. Dawson, D. Herschlag, Evaluating the Potential for Halogen Bonding in the Oxyanion Hole of Ketosteroid Isomerase Using Unnatural Amino Acid Mutagenesis, *Acs Chemical Biology*, 4 (2009) 269-273.
- [15] L.A. Hardegger, B. Kuhn, B. Spinnler, L. Anselm, R. Ecabert, M. Stihle, B. Gsell, R. Thoma, J. Diez, J. Benz, J.-M. Plancher, G. Hartmann, D.W. Banner, W. Haap, F. Diederich, Systematic Investigation of Halogen Bonding in Protein-Ligand Interactions, *Angewandte Chemie-International Edition*, 50 (2011) 314-318.
- [16] R.G. Eckenhoff, J.S. Johansson, Molecular interactions between inhaled anesthetics and proteins, *Pharmacological Reviews*, 49 (1997) 343-367.
- [17] R.Y. Liu, P.J. Loll, R.G. Eckenhoff, Structural basis for high-affinity volatile anesthetic binding in a natural 4-helix bundle protein, *Faseb Journal*, 19 (2005) 567-576.
- [18] S.M. Pop, N. Gupta, A.S. Raza, S.W. Ragsdale, Transcriptional activation of dehalorespiration - Identification of redox-active cysteines regulating dimerization and DNA binding, *Journal of Biological Chemistry*, 281 (2006) 26382-26390.
- [19] A. Memic, M.R. Spaller, How Do Halogen Substituents Contribute to Protein-Binding Interactions? A Thermodynamic Study of Peptide Ligands with Diverse Aryl Halides, *Chembiochem*, 9 (2008) 2793-2795.
- [20] P. Zien, M. Bretner, K. Zastapilo, R. Szyszka, D. Shugar, Selectivity of 4,5,6,7-tetrabromobenzimidazole as an ATP-competitive potent inhibitor of protein kinase CK2 from various sources, *Biochemical and biophysical research communications*, 306 (2003) 129-133.

- [21] S. Sarno, H. Reddy, F. Meggio, M. Ruzzene, S.P. Davies, A. Donella-Deana, D. Shugar, L.A. Pinna, Selectivity of 4,5,6,7-tetrabromobenzotriazole, an ATP site-directed inhibitor of protein kinase CK2 ('casein kinase-2'), *FEBS letters*, 496 (2001) 44-48.
- [22] M.A. Pagano, J. Bain, Z. Kazimierczuk, S. Sarno, M. Ruzzene, G. Di Maira, M. Elliott, A. Orzeszko, G. Cozza, F. Meggio, L.A. Pinna, The selectivity of inhibitors of protein kinase CK2: an update, *Biochemical Journal*, 415 (2008) 353-365.
- [23] R. Battistutta, M. Mazzorana, L. Cendron, A. Bortolato, S. Sarno, Z. Kazimierczuk, G. Zanotti, S. Moro, L.A. Pinna, The ATP-binding site of protein kinase CK2 holds a positive electrostatic area and conserved water molecules, *Chembiochem*, 8 (2007) 1804-1809.
- [24] S. Sarno, E. Papinutto, C. Franchin, J. Bain, M. Elliott, F. Meggio, Z. Kazimierczuk, A. Orzeszko, G. Zanotti, R. Battistutta, L.A. Pinna, ATP Site-Directed Inhibitors of Protein Kinase CK2: An Update, *Current Topics in Medicinal Chemistry*, 11 (2011) 1340-1351.
- [25] P. Zien, J.S. Duncan, J. Skierski, M. Bretner, D.W. Litchfield, D. Shugar, Tetrabromobenzotriazole (TBBt) and tetrabromobenzimidazole (TBBz) as selective inhibitors of protein kinase CK2: Evaluation of their effects on cells and different molecular forms of human CK2, *Biochimica Et Biophysica Acta-Proteins and Proteomics*, 1754 (2005) 271-280.
- [26] R. Wasik, M. Lebska, K. Felczak, J. Poznanski, D. Shugar, Relative role of halogen bonds and hydrophobic interactions in inhibition of human protein kinase CK2 $\alpha$  by tetrabromobenzotriazole and some C(5)-substituted analogues, *J. Phys. Chem. B*, 114 (2010) 10601-10611.
- [27] R. Wasik, P. Winska, J. Poznanski, D. Shugar, Synthesis and Physico-Chemical Properties in Aqueous Medium of All Possible Isomeric Bromo Analogues of Benzo-1H-Triazole, Potential Inhibitors of Protein Kinases, *The journal of physical chemistry*, 116 (2012) 7259-7268.
- [28] R. Wasik, P. Winska, J. Poznanski, D. Shugar, Isomeric Mono-, Di-, and Tri-Bromobenzo-1H-Triazoles as Inhibitors of Human Protein Kinase CK2 $\alpha$ , *PloS one*, 7 (2012) e48898-e48898.
- [29] J. Poznanski, A. Najda, M. Bretner, D. Shugar, Experimental ( $^{13}\text{C}$  NMR) and theoretical (ab initio molecular orbital calculations) studies on the prototropic tautomerism of benzotriazole and some derivatives symmetrically substituted on the benzene ring, *J. Phys. Chem. A*, 111 (2007) 6501-6509.
- [30] <http://www.rcb.org>, downloaded may 2012.
- [31] E. Krieger, K. Joo, J. Lee, J. Lee, S. Raman, J. Thompson, M. Tyka, D. Baker, K. Karplus, Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8, *Proteins*, 77 Suppl 9 (2009) 114-122.
- [32] D. Hauchecorne, N. Nagels, B.J. van der Veken, W.A. Herrebout, C-X... $\pi$  halogen and C-H... $\pi$  hydrogen bonding: interactions of CF<sub>3</sub>X (X = Cl, Br, I or H) with ethene and propene, *Physical Chemistry Chemical Physics*, 14 (2012) 681-690.
- [33] A.S. Konagurthu, J.C. Whisstock, P.J. Stuckey, A.M. Lesk, MUSTANG: A multiple structural alignment algorithm, *Proteins-Structure Function and Bioinformatics*, 64 (2006) 559-574.
- [34] D.W. Marquardt, AN ALGORITHM FOR LEAST-SQUARES ESTIMATION OF NONLINEAR PARAMETERS, *Journal of the Society for Industrial and Applied Mathematics*, 11 (1963) 431-441.
- [35] <http://www.gnuplot.info>, version 4.6.
- [36] M.A.A. Ibrahim, Performance Assessment of Semiempirical Molecular Orbital Methods in Describing Halogen Bonding: Quantum Mechanical and Quantum Mechanical/Molecular Mechanical-Molecular Dynamics Study, *Journal of Chemical Information and Modeling*, 51 (2011) 2549-2559.
- [37] M.A.A. Ibrahim, AMBER Empirical Potential Describes the Geometry and Energy of Noncovalent Halogen Interactions Better than Advanced Semiempirical Quantum Mechanical Method PM6-DH2X, *Journal of Physical Chemistry B*, 116 (2012) 3659-3669.
- [38] M. Kolar, K. Berka, P. Jurecka, P. Hobza, On the Reliability of the AMBER Force Field and its Empirical Dispersion Contribution for the Description of Noncovalent Complexes, *Chemphyschem*, 11 (2010) 2399-2408.

- [39] W.L. Jorgensen, P. Schyman, Treatment of Halogen Bonding in the OPLS-AA Force Field: Application to Potent Anti-HIV Agents, *Journal of Chemical Theory and Computation*, 8 (2012) 3895-3901.
- [40] Y. Lu, H. Li, X. Zhu, W. Zhu, H. Liu, How Does Halogen Bonding Behave in Solution? A Theoretical Study Using Implicit Solvation Model, *Journal of Physical Chemistry A*, 115 (2011) 4467-4475.
- [41] E. De Moliner, N.R. Brown, L.N. Johnson, Alternative binding modes of an inhibitor to two different kinases, *European Journal of Biochemistry*, 270 (2003) 3174-3181.
- [42] R. Battistutta, M. Mazzorana, S. Sarno, Z. Kazimierczuk, G. Zanotti, L.A. Pinna, Inspecting the structure-activity relationship of protein kinase CK2 inhibitors derived from tetrabromo-benzimidazole, *Chemistry & biology*, 12 (2005) 1211-1219.
- [43] T.M. Beale, M.G. Chudzinski, M.G. Sarwar, M.S. Taylor, Halogen bonding in solution: thermodynamics and applications, *Chem. Soc. Rev.*, (2012).
- [44] R. Battistutta, E. De Moliner, S. Sarno, G. Zanotti, L.A. Pinna, Structural features underlying selective inhibition of protein kinase CK2 by ATP site-directed tetrabromo-2-benzotriazole, *Protein Science*, 10 (2001) 2200-2206.
- [45] A.D. Ferguson, P.R. Sheth, A.D. Basso, S. Paliwal, K. Gray, T.O. Fischmann, H.V. Le, Structural basis of CX-4945 binding to human protein kinase CK2, *FEBS letters*, 585 (2011) 104-110.
- [46] Z. Xu, Z. Liu, T. Chen, T. Chen, Z. Wang, G. Tian, J. Shi, X. Wang, Y. Lu, X. Yan, G. Wang, H. Jiang, K. Chen, S. Wang, Y. Xu, J. Shen, W. Zhu, Utilization of Halogen Bond in Lead Optimization: A Case Study of Rational Design of Potent Phosphodiesterase Type 5 (PDE5) Inhibitors, *Journal of medicinal chemistry*, 54 (2011) 5607-5611.