J Comput Aided Mol Des (2012) 26:935–945
DOI 10.1007/s10822-012-9592-8

Using halogen bonds to address the protein backbone: a systematic evaluation

Rainer Wilcken · Markus O. Zimmermann · Andreas Lange · Stefan Zahn · Frank M. Boeckler

Received: 18 June 2012 / Accepted: 19 July 2012 / Published online: 4 August 2012

Abstract Halogen bonds are specific embodiments of the sigma hole bonding paradigm. They represent directional interactions between the halogens chlorine, bromine, or iodine and an electron donor as binding partner. Using quantum chemical calculations at the MP2 level, we systematically explore how they can be used in molecular design to address the omnipresent carbonyls of the protein backbone. We characterize energetics and directionality and elucidate their spatial variability in sub-optimal geometries that are expected to occur in protein–ligand complexes featuring a multitude of concomitant interactions. By deriving simple rules, we aid medicinal chemists and chemical biologists in easily exploiting them for scaffold decoration and design. Our work shows that carbonyl–halogen bonds may be used to expand the patentable medicinal chemistry space, redefining halogens as key features. Furthermore, this data will be useful for implementing halogen bonds into pharmacophore models or scoring functions making the QM information available for automatic molecular recognition in virtual high throughput screening.

Keywords Sigma hole · Halogen bonding · Molecular recognition · Protein backbone · QM-based interaction geometries · Medicinal chemistry

R. Wilcken · M. O. Zimmermann · A. Lange · F. M. Boeckler (✉)
Laboratory for Molecular Design and Pharmaceutical Biophysics, Department of Pharmaceutical and Medicinal Chemistry, Institute of Pharmacy, Eberhard-Karls-University Tübingen, Auf der Morgenstelle 8, 72076 Tübingen, Germany
e-mail: frank.boeckler@uni-tuebingen.de

R. Wilcken
MRC Laboratory of Molecular Biology, Hills Rd, Cambridge CB2 0QH, UK

S. Zahn
Wilhelm-Ostwald-Institut für Physikalische und Theoretische Chemie, Universität Leipzig, Linnéstr. 2, 04103 Leipzig, Germany

Introduction

Halogen bonding can be described as an attractive interaction of the type R–X…D–R', where X represents chlorine, bromine, or iodine, and D can be any kind of Lewis base. The interaction has been broadly recognized in materials sciences since the 1970s [1, 2], but its occurrence in biological systems has been studied only recently, e.g. through PDB evaluations [3]. The driving force of the interaction can be explained within the σ-hole concept—the halogen atom possesses a characteristic crown of positive charge due to a deficiency in electron density in the outer lobe of the p_z orbital (where z is chosen as the R–X axis) [4–8].

Aromatic halogenated ligands carrying chloro, bromo or iodo substituents can theoretically form halogen bonds with any electron donor. A recent PDB-evaluation of halogen bonds in protein–ligand complexes revealed that the one involving the carbonyl function of the protein backbone is the most frequent [9]. As for the halogen atoms, chlorine is by far the most common substituent, followed by bromine and iodine. In the process of ligand de novo-design and scaffold decoration, halogen bonds are generally rather neglected. Up to now, most halogen bonds found in crystal structures were typically not designed, but rather found serendipitously [10, 11]. Considering that every amino acid at least carries one carbonyl function, the

probability of finding an addressable carbonyl oxygen in a binding site is extremely high. Furthermore, statistical analyses of known protein–ligand complexes do not provide explicit information about the quality of the halogen bond and the frequency of occurrence does not necessarily correlate with the strength of this interaction. The frequency of scaffold decoration with chlorine, bromine, and iodine will also be reflected in a purely statistical assessment.

Only recently, Hardegger et al. [12] have studied halogen bonding systematically in a series of covalent human Cathepsin L inhibitors, demonstrating that halogen bonds are valuable interaction types for drug discovery. Other recently published studies of halogen bonding in lead optimization have cumulated experimental evidence for the broad applicability of halogen bonding in medicinal and biological chemistry, addressing PDE5 [13], PIM1 kinase [10], and MEK1 kinase [14]. However, the use of halogen bonding is not only limited to lead optimization, which aims at enhancing ligand efficacy, selectivity, and ADMET properties by decorating and transforming an already established chemical scaffold, e.g. by introducing halogen atoms. It can also be more generally applied to lead identification, facilitating to establish binding modes based on halogen bonding as core interaction. We have devised a strategy for lead discovery by halogen-enriched fragment libraries (HEFLibs), which aims at enhancing the statistical occurence of halogen bonds in early screening [15]. Moreover, HEFLibs automatically offer a multitude of synthetic cross-coupling mechanisms for extending the new lead structure. We have demonstrated this by identifying and optimizing new stabilizers of the p53 mutant Y220C [15], which is rapidly unfolded at body temperature based on the thermal destabilization caused by this mutation [16, 17].

Some quantum chemical studies using small model systems have been performed to gain insights into the nature and strength of halogen bonds [18–20]. QM/MM calculations on protein–ligand complexes have also been conducted [9, 21] and similarities as well as differences between halogen bonding and hydrogen bonding have been discussed [22–25].

As a basis for inspiring more rational and computer-aided drug design utilizing this type of interaction, we have comprehensively evaluated optimal and non-optimal halogen bond geometries between aromatic halogenated ligands and the carbonyl moiety of the protein backbone using small model systems. Recently, we have addressed halogen–sulfur contacts toward methionine using a similar methodology [20]. We aim at helping to understand the energetical dependencies of this interaction by visualization of these results in protein binding sites. We have proven applicability of this theoretical study, having already used it as a knowledge base for developing HEFLibs for lead discovery [15].

Choice of model systems

Model systems for use in quantum chemical calculations have to be chosen carefully. A suitable model system to represent the backbone carbonyl moiety of a protein should be small, to allow high-level quantum chemical calculations to be conducted, but should include all relevant atoms. At the same time it is crucial to reflect the biological and chemical nature of the protein substructure as correctly as possible.

Several model systems differing in size were considered and preliminary calculations were performed. Riley et al. [19] gave interesting insights into the C–Br…O bond using acetone and bromobenzene as model systems. The smallest biosimilar system to the protein backbone is N-methylacetamide, which as a capped peptide bond represents the protein backbone in our study (hence abbreviated "BB" herein). The small size of this model system allows high quality calculations with larger basis sets, while reducing the risk of unwanted secondary interactions besides the halogen bond.

Iodobenzene, bromobenzene and chlorobenzene are employed as ligand models. Substituted aromatic or heteroaromatic scaffolds are common building blocks in standard medicinal chemistry and they can be decorated by halogenation for lead optimization purposes. More reactive species like the smaller alkyl halides are not chosen as they are unlikely to be part of any design approach aiming to develop strong but reversible binders. In order to put the halogen bond strengths into context, we also employ benzene and phenol as ligand models, representing weak and moderately strong hydrogen bonding to the backbone carbonyl moiety.

Computational details

Structure optimizations

Geometry optimizations were carried out at the MP2/TZVPP [26] level using the TURBOMOLE 6.2 suite of programs [27, 28]. Relativistic effects for iodine were considered by an effective core potential (ECP) [29]. The calculations were done in combination with the resolution of identity (RI) technique [30–33] and the frozen core approximation. The frozen core orbitals were attributed by the default setting in TURBOMOLE by which all orbitals possessing energies below 3.0 au are considered core orbitals. Using the frozen core approximation for treating

relativity in iodine may have effects on the polarizability of the outer electrons and consequently might not represent dispersion correctly. The SCF convergence criterion was increased to 10^{-8} Hartree for all calculations. Interaction energies were counterpoise-corrected using the procedure of Boys and Bernardi [34] where indicated in order to correct for basis set superposition errors (BSSEs).

CCSD(T) calculations

Structure optimizations for iodobenzene, the backbone model system (BB) and the iodobenzene–BB complex were performed with the TURBOMOLE suite employing the SCS-MP2 [35] method and the def2-QZVPP [26] basis set. Additionally, we performed MP2 single point calculations with the cc-pVTZ and cc-pVQZ [36–38] basis sets. These MP2 calculations were done in combination with the RI technique [30–33, 39] and the frozen core approximation as described above. Relativistic effects for iodine were considered by an effective core potential (ECP) [38].

CCSD(T) calculations were carried out with the cc-pVTZ [36–38] basis set employing the MOLPRO [40] program package. Relativistic effects for iodine were again considered by a relativistic pseudopotential. All CCSD(T) and MP2 energies were counterpoise-corrected using the procedure of Boys and Bernardi [34]. The contribution of higher order correlation energy was determined by following scheme [41, 42]:

$$\Delta E_{CBS}^{CCSD(T)} = \Delta E_{CBS}^{MP2} + (\Delta E_{cc-pVTZ}^{CCSD(T)} - \Delta E_{cc-pVTZ}^{MP2})$$

This is based on the assumption that the difference between the CCSD(T) and MP2 interaction energies ($\Delta E^{CCSD(T)} - \Delta E^{MP2}$) depends only slightly on basis set size and therefore can be determined with small or medium-sized basis sets like cc-pVTZ. ΔE_{CBS}^{MP2} is the MP2 energy at the complete basis set limit which was obtained by the extrapolation procedure proposed by Halkier et al. [43]:

$$\Delta E_{CBS}^{MP2} = \frac{\Delta E_X^{MP2} X^3 - \Delta E_Y^{MP2} Y^3}{X^3 - Y^3}$$

where X and Y are the cardinal numbers of the cc-pVTZ and cc-pVQZ basis set, respectively.

Distance scans

All distance scans were performed using the optimized MP2/TZVPP geometries as starting points. The bond distances $d_{X\ldots O}$ ($d_{H\ldots O}$ for benzene) were then elongated or shortened in 2 pm steps (5 pm further away from the minimum), with the rest of the ligand structure (halobenzene or benzene) transformed accordingly. For each step, single-point (SCF) calculations were performed using MP2/TZVPP.

Spherical scans

Input files were generated from optimized MP2/TZVPP geometries subject to certain constraints. The C–X group was placed on the vector defined by the C=O group, with both the angle $\alpha_{C4\text{-}O5\text{-}X6}$ and the angle $\alpha_{O5\text{-}X6\text{-}C7}$ constrained to 180°; the bond distance was freely optimized. In order to generate a full sphere of input geometries for subsequent calculations, the optimized structure was transformed as follows: The oxygen atom was placed on the origin of the coordinate system and the entire complex was rotated until the halogen atom was positioned on the positive x-axis. Let α denote the angle of rotation counter-clockwise around the z-axis, and β denote the angle of rotation counter-clockwise around the x-axis. α was gradually increased from 0° to 180° in steps of 5°. For each α-value, β was varied from 0° to 355° in steps of 5°, leading to a total number of 2,664 halogen positions distributed on a sphere. The structure of the ligand was not altered during this transformation process. Calculations were done as single point (SCF) calculations using the MP2/TZVPP method.

For visualization purposes several Python scripts were written and executed in PyMOL v1.2 [44]. Interaction energies were partitioned into bins and spectrum colors (red to blue to purple) were assigned. At the positions of the halogens small CONE objects with the appropriate coloring were generated.

Results

Energetics of backbone halogen bonding at MP2 and CCSD(T) level

Performing free geometry optimizations at the MP2/TZVPP level, halogen-bonded complexes for all three halobenzenes and hydrogen-bonded complexes for benzene and phenol are obtained (Fig. 1). Interaction energies are calculated as adduct formation of the complexes from isolated ligand models with N-methylacetamide (Table 1; Scheme 1).

Among the halogen-bonded complexes, the strongest interaction is observed for the iodobenzene…BB complex with −14.2 kJ/mol, while bromobenzene and chlorobenzene complexes show weaker interaction energies (−9.0 kJ/mol and -5.6 kJ/mol, respectively). This order of halogen bond strengths (Cl < Br < I) has been reported previously [8, 22], regardless of the Lewis base interaction partner, and is consistent with the magnitude of the σ-hole on the halogen of the model systems (Fig. 2).

All halogen bonds show X…O distances below the sum of van der Waals radii of the two atoms. Despite the size of

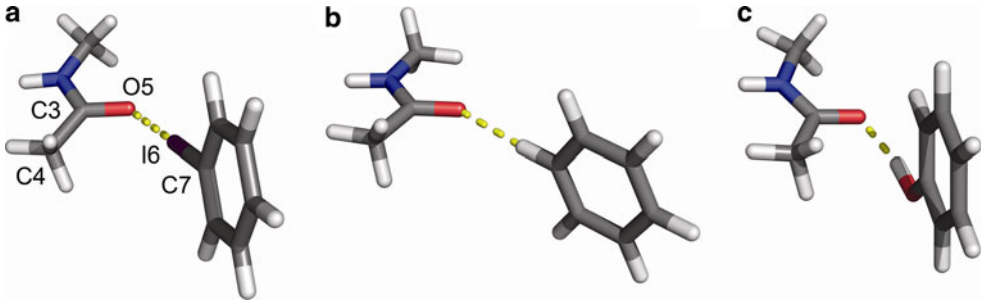


Fig. 1 Optimized geometries of the **a** halobenzene…BB, **b** benzene…BB, and **c** phenol…BB complexes

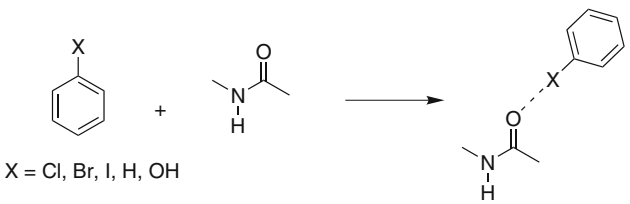
Table 1 Interaction energies (in kJ/mol) for halogen-bonded model complexes

Complex	ΔE [kJ/mol]	Method	$d_{O5\text{-}X6/H6}$ (%vdW)[a]	$\alpha_{C3\text{-}O5\text{-}X6/H6}$	$\delta_{C4\text{-}C3\text{-}O5\text{-}X6/H6}$	$\alpha_{O5\text{-}X6/H6\text{-}C7}$
C_6H_5I…BB	−14.2 (−18.8)	MP2/TZVPP	302 pm (86.3%)	118.7°	−30.3°	175.6°
C_6H_5Br…BB	−9.0 (−11.8)	MP2/TZVPP	304 pm (90.2%)	116.7°	−30.7°	177.4°
C_6H_5Cl…BB	−5.6 (−8.2)	MP2/TZVPP	312 pm (95.4%)	106.7°	−39.5°	171.2°
C_6H_6…BB	−8.4	MP2/TZVPP	230 pm (84.6%)	107.6°	178.7°	175.8°
C_6H_5OH…BB	−44.5	MP2/TZVPP	174 pm (64.4%)	114.3°	3.1°	167.8°
C_6H_5I…BB	−17.6	CCSD(T)/CBS[b]	306 pm (87.4%)	119.4°	−31.3°	175.2°

Interaction energies for complex formation with benzene and phenol (H…O contacts) are given for comparison. Energies were corrected for BSSE using the Counterpoise correction (uncorrected energies are given in brackets)

[a] Percentage of the sum of the van der Waals radii of the two atoms directly involved in bonding

[b] CCSD(T) calculations were performed on the SCS-MP2/QZVPP minimum geometry and a basis set extrapolation scheme was employed (see methods)



Scheme 1 Adduct formation reaction for the halobenzene…BB, benzene…BB, and phenol…BB complexes

iodine, the I…O bond is the shortest, implying a rather favorable interaction. Interestingly, the equilibrium distances for all three halogen bonds are almost identical in range (3.0–3.1 Å). The MP2 complex formation energy of benzene…BB is −8.4 kJ/mol, which is between the values obtained for bromobenzene…BB and chlorobenzene…BB. The H…O (230 pm) is much shorter than the halogen bond distances. The largest interaction energy is obtained for the hydrogen bond involving phenol with −44.5 kJ/mol. The proximity of the terminal methyl group of N-methylacetamide to the aromatic ring of the phenol indicates the presence of an additional CH…π contact, making the overall complex formation energy even more favorable.

To describe the strength of an iodine–oxygen halogen bond with chemical accuracy, we perform CCSD(T) calculations employing a basis set extrapolation scheme (see methods) and obtain an interaction energy of −17.6 kJ/mol. We find that the MP2/TZVPP method may slightly underestimate the quality of the interaction. In similar calculations of halogen-sulfur contacts, we have observed that larger interaction energies for all halogen-bonded systems are obtained when switching from a triple-ζ basis set (TZVPP) to the quadruple-ζ basis QZVPP [20]. Considering the CBS extrapolation used in the CCSD(T) reference calculation, we assume that part of the observed underestimation is due to the smaller basis set, but might be

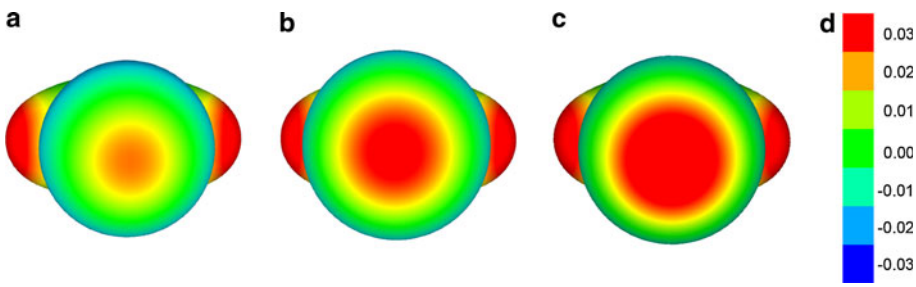


Fig. 2 Electrostatic potentials plotted onto the isodensity surfaces at 0.003 au for **a** chlorobenzene, **b** bromobenzene and **c** iodobenzene. The halogen is facing the viewer in each case. *Color ranges* of energies in atomic units are shown in (**d**)

caused as well by a shortcoming of the Counterpoise correction approach. MP2 without Counterpoise correction (Table 1) gives a complex formation energy of −18.8 kJ/mol for iodobenzene…BB. Thus, uncorrected MP2/TZVPP, which is used throughout the rest of the study, can be considered a reasonable approximation to higher level calculations.

Distance dependencies of halogen bonds

Ligand binding is based on the interplay of several competing primary and secondary interactions, such as hydrogen or halogen bonds. As a consequence, optimal geometries of these interactions are not always observed. In the following sections, we will elucidate the energetical behaviour of the halobenzene…BB complexes in non-ideal bond distances and angles. Such data is crucial for parameterization of non-QM-based in silico techniques for recognizing halogen bonding.

Introducing a halogen atom (X=Cl, Br, I) into a scaffold should increase its affinity towards its target protein. By default, a C–H bond within the aromatic moiety will be replaced by C–X. Consequently, the contribution of an H…O contact to binding affinity has to be compared with an X…O contact at the same position in the scaffold. Thus, we evaluate the complex formation energy of the halobenzene…BB complexes relative to those of benzene…BB at different bond distances. The ideal bond distances and the minimum energies for the benzene and halobenzene complexes are listed in Table 2. For comparative reasons, the energies at the equilibrium distance of iodine is listed as well. As discussed before, the optimal distances for the halobenzenes are very similar between 302 and 312 pm. The complete distance scans are shown in Fig. 3. Resulting from a lower optimal distance, the curve for benzene is shifted to the left. For the purpose of scaffold decoration, its intersection points with the curves of the halobenzenes can be assumed as an energetical cutoff for the exchange of a hydrogen with a halogen atom (Table 2, right column).

Due to repulsive forces, the curves become rather steep when reducing the X/H…O distance (starting from the optimal bond distance). An increase in distance leads to a slow convergence toward a complex formation energy of zero.

Table 2 Comparison of MP2/TZVPP complex formation energies (in kJ/mol) at optimal distances and at the equilibrium distance of iodobenzene…BB

Complex	$d_{O5-X6/H6}$ (pm)	$\Delta E_{MP2/TZVPP}$ at optimal distances	$\Delta E_{MP2/TZVPP}$ at 302 pm	Intersection point with benzene
C_6H_5I…BB	302	−18.8	−18.8	≈263 pm
C_6H_5Br…BB	304	−11.8	−11.8	≈274 pm
C_6H_5Cl…BB	312	−8.2	−8.0	≈292 pm
C_6H_6…BB	230	−10.6	−6.8	N/A

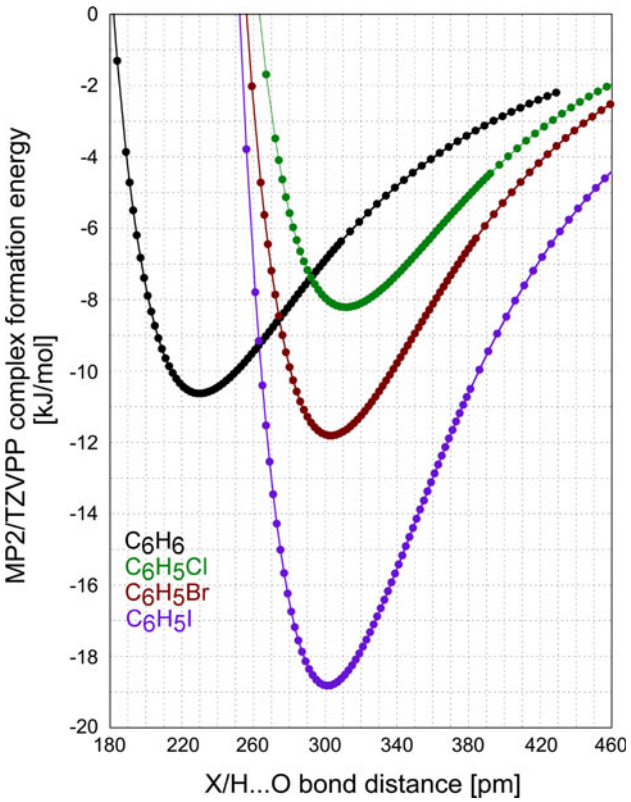


Fig. 3 Distance scan plots for the three halobenzene…BB complexes and benzene…BB as comparison. Plot of X…O (H…O) distance versus complex formation energy for iodobenzene, bromobenzene, chlorobenzene and benzene complexes with BB (MP2/TZVPP)

Spatial behaviour of halogen bonds: sigma-hole directionality

We have investigated the energetic impact of σ-hole bond directionality. It is generally agreed that the driving force of halogen bonding is the σ-hole, a crown of positive charge on chlorine, bromine and iodine atoms on the opposite side of its covalent bond (Fig. 2). The σ-hole concept explains halogen bonding as an electrostatic interaction similar to hydrogen bonding [4, 22]. One study has suggested that halogen bonds may be driven by both electrostatic and dispersive forces, using symmetry-adapted perturbation theory [18]. Very recently, this method was applied to study the tunability of halogen bonds by fluorine [45]. Decomposition of interaction energies into dispersive and electrostatic/polarization contributions suggests that at equilibrium distance both are similar in magnitude. The higher the most positive electrostatic potential of the σ-hole, the more the electrostatic character of the interaction is strengthened, while the dispersive character is simultaneously weakened based on decreased polarizability of the halogen.

While a purely electrostatics-driven interaction would translate into a very directional bond with strong penalties for deviations from the ideal bond angle $\alpha_{O5\text{-}X6\text{-}C7}$, a dispersion-driven interaction could tolerate such deviations to a larger extent. Still it should be noted that also polarizability of halogens are most likely rather anisotropic.

Preferential directionality certainly has great impact in molecular design approaches, e.g. for scaffold placement. To evaluate the energy penalties of non-ideal σ-hole bond angles, the halobenzene…BB complexes were optimized with the bond angles $\alpha_{C3\text{-}O5\text{-}X6}$ and $\alpha_{O5\text{-}X6\text{-}C7}$ constrained to 179.98°. This is done in order to rule out any secondary interactions of the aromatic ring with the methyl moieties of the BB system. The bond angle $\alpha_{O5\text{-}X6\text{-}C7}$ is then successively varied by 5° in both directions and MP2/TZVPP single point calculations are performed. The results of the scans for all three halobenzenes are shown in Fig. 4. The scans clearly demonstrate that there is a strong penalty for deviating from the optimal σ-hole angle which is usually close to 180°. We find that deviations by 30° lead to a decrease in interaction energy of ~50–60 %. This behavior is in accordance with other works [4, 6, 8, 22, 24] and suggests that halogen bonding is largely driven by electrostatic forces, while dispersion might play a role in cases of less positive σ-holes and highly anisotropic polarizabilities.

Spatial behavior of halogen bonds: spherical scans

The next logical step towards integrating halogen bonds in molecular design tools is to evaluate their spatial behavior by setting up a spherical scan around the carbonyl oxygen. While keeping the halobenzene at equilibrium distance at all times and varying both dihedral angles δ_1 (out-of-plane rotation) and δ_2 (in-plane rotation) that define the interaction geometry successively by 5° (Fig. 5) we obtained a full sphere of 2,664 variations of the optimized input geometry. As before, the optimized input geometry was obtained by constraining the angles $\alpha_{C3\text{-}O5\text{-}X6}$ and $\alpha_{O5\text{-}X6\text{-}C7}$ to 180°. This was done to avoid unsymmetrical artifacts due to bias in the input structure.

The obtained interaction spheres for iodobenzene…BB from different perspectives are presented in Fig. 6. Each individual colored region represents one interaction geometry and its corresponding energy. The most favorable areas of interaction with energies ≤ -16.5 kJ/mol surround the carbonyl oxygen in an elliptical manner, as can be seen in the lower images of Fig. 6a or b. Towards the outer borders energy decreases fast and significantly due to van der Waals clashes of iodobenzene with BB. The energy of the constrained (180°) input geometry deviates by around 3.6 kJ/mol from the optimum. In general, there is a very large area of favorable interaction geometries, which makes the carbonyl oxygen easily addressable within a protein's binding pocket.

For bromobenzene and chlorobenzene the distribution of the favorable interaction areas is very similar to iodobenzene apart from a substantial overall decrease in interaction energy (see Fig. 7). The best complex formation energies for bromobenzene lie between −10.0 and −11.8 kJ/mol (Fig. 7a), for chlorobenzene between −6.5 and −8.2 kJ/mol (Fig. 7b).

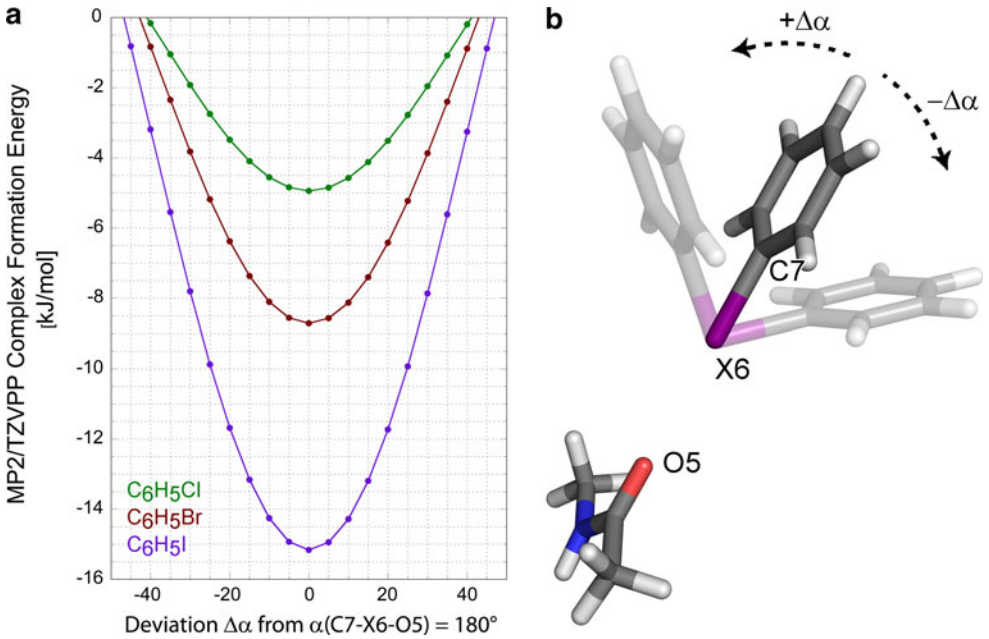


Fig. 4 **a** Impact of deviations from 180° in the σ-hole angle $\alpha_{O5\text{-}X6\text{-}C7}$ for halobenzene…BB complexes. Energies were calculated using MP2/TZVPP. **b** Illustration of the scan setup. Deviations ($\pm\Delta\alpha$) from optimal α are depicted as faded

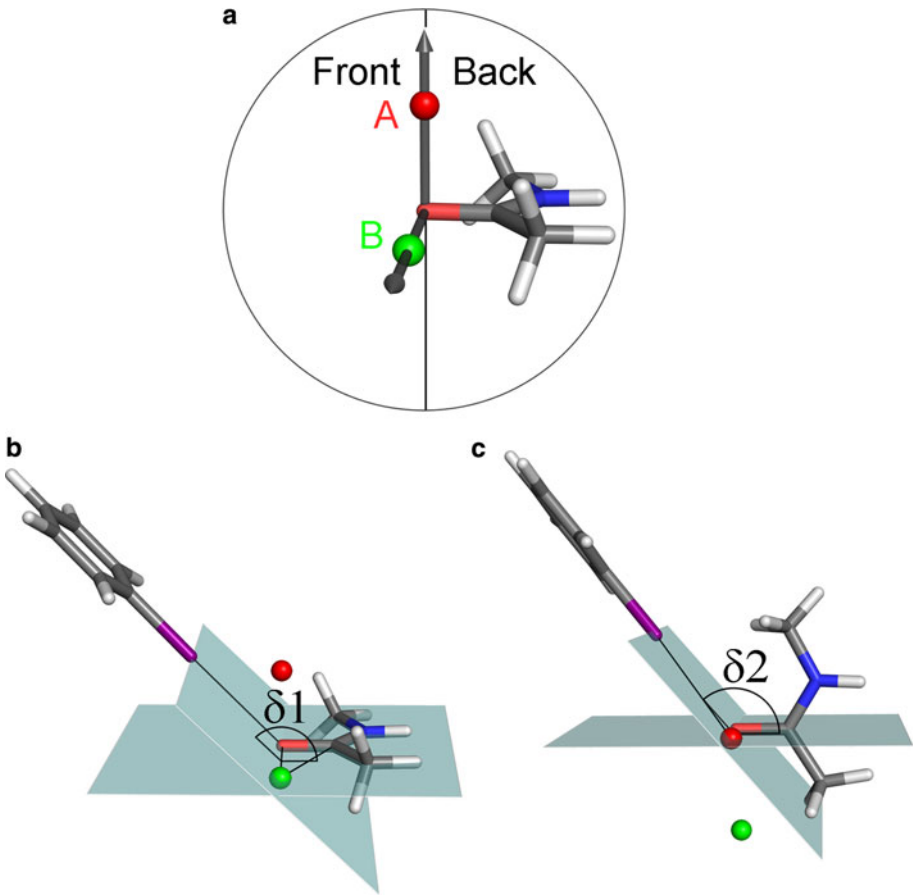


Fig. 5 Definitions of spherical scans for halobenzene…BB, starting from the constrained input geometry. **a** Definition of dummy points A and B used to define dihedral angles. **b** Illustration of dihedral angle δ_1 (out-of-plane rotation). δ_1 represents the rotation around an axis through dummy point B and atom O5, which is part of the N–C=O plane and is oriented perpendicularly to the plane defined by C3, O5, and dummy point A. **c** Illustration of dihedral angle δ_2 (in-plane rotation). δ_2 represents the rotation around an axis from dummy point A to atom O5, which is oriented perpendicularly to the N–C=O plane

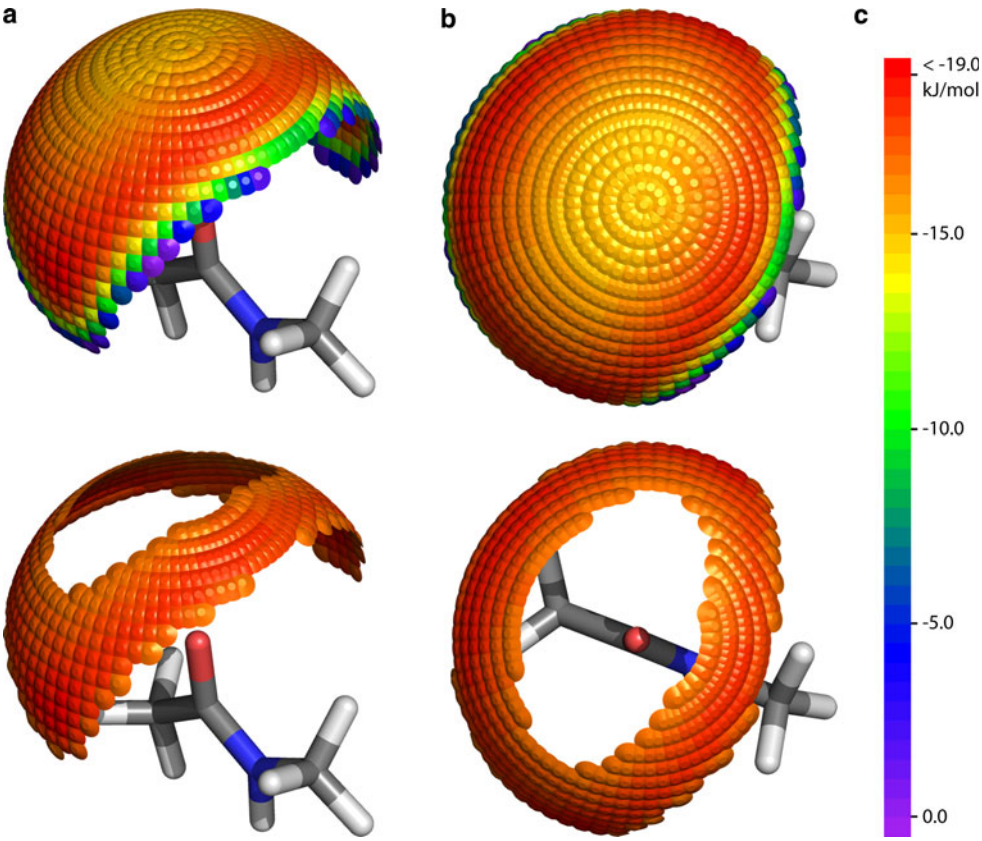


Fig. 6 MP2/TZVPP spherical scans for the iodobenzene…BB complex, starting from the constrained 180°-input geometry. **a** Spherical scan for iodobenzene from the side. Favorable areas with energies <-16.5 kJ/mol are shown below. **b** Spherical scan for iodobenzene from the top. Favorable areas with energies <-16.5 kJ/mol are shown below. **c** Spectrum colors and associated energies in kJ/mol

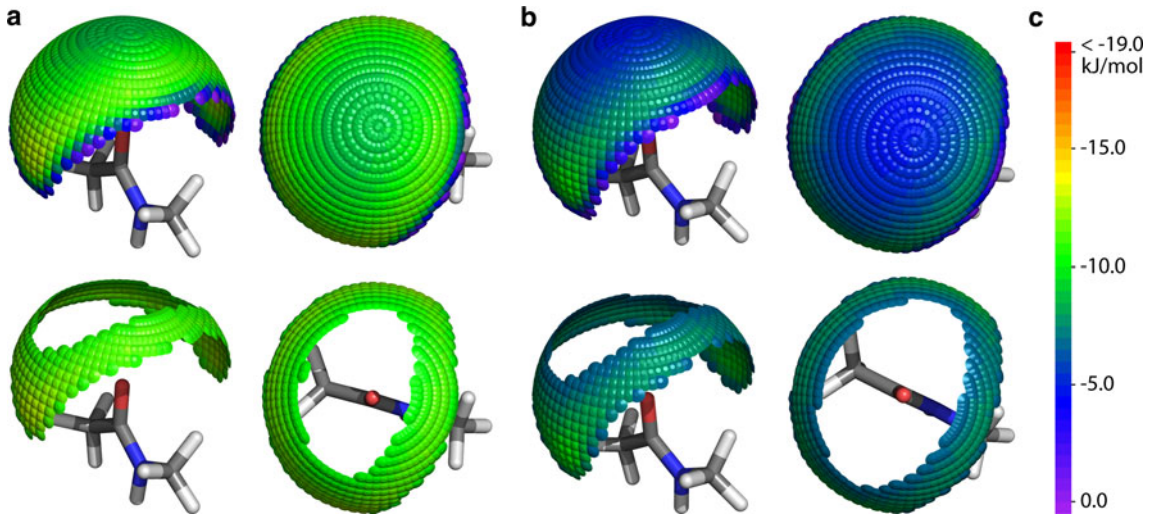


Fig. 7 MP2/TZVPP spherical scans for bromobenzene…BB and chlorobenzene…BB, starting from the constrained 180°-input geometries. **a** Spherical scans for bromobenzene from the side and from the top perspective. Favorable areas with energies <−10.0 kJ/mol are shown below. **b** Spherical scans for chlorobenzene from the side and from the top perspective. Favorable areas with energies <−6.5 kJ/mol are shown below. **c** Spectrum colors and associated energies in kJ/mol

Discussion

As described previously in other works, we confirm that the halogen bond strength increases in the order Cl < Br < I. The iodobenzene halogen bond with our backbone model system is notably stronger than the weak hydrogen bonds between benzene and the carbonyl oxygen. The equilibrium distances for all halobenzene…BB complexes are very similar in range, indicating mutual interchangeability of the halogens, despite their differences in atom radius.

The comparison of distance scans between benzene and the three halobenzenes gives a first indication of when an introduction of a halogen moiety into an unsubstituted aromatic scaffold becomes favorable for ligand affinity. σ-hole scans show that the halogen bond is highly directional and even small deviations from the ideal angle lead to severe losses in interaction energy.

As we have recently shown, the favorable spherical angles of a halogen bond toward the sulfur atom of methionine are rather restricted [20]. With respect to their interaction energies, only few geometries lie within a small margin of the optimum (regardless of σ-hole angle or bond distance). In contrast, halogen bonds toward the carbonyl oxygen can exist in a large variety of favorable spherical orientations. Despite the limitations of the σ-hole angle or the bond distance discussed above, the spherical degrees of freedom make this moiety a lot easier to address within any binding pocket. In addition, backbone carbonyls are ubiquitiously available and often solvent-exposed within the binding site. Most interestingly, the polarization-based induction of σ-holes was very recently reported by Hennemann et al. [46]. As a consequence the local binding site environment may be important for tuning the quality of the halogen bonds.

But how do these results compare with experimentally obtained crystal structures? Projecting the computed interaction spheres onto carbonyl atoms in several PDB structures shows that there are indeed a number of structures that fit the model very well (Fig. 8).

In all of these examples the σ-hole angle lies between 165° and 176°, which is close to the optimum (see Table 1, right column). It becomes clear that the halogen–oxygen contact lies within the energetically most favorable areas of the sphere, as seen in Figs. 6 and 7. For the alignment and orientation of the spheres, the BB moiety of the model system is mapped onto the backbone peptide bond of the residue of interest (minimizing the RMSD of the heavy atoms). This is an easy and fast way to determine the quality of an existing halogen bond in a protein–ligand complex. For the purpose of scaffold decoration it is possible to map these spheres onto any carbonyl moiety in the binding pocket to check for energetically favorable halogen–oxygen contacts. In doing so, we have recently designed halogen-enriched fragment libraries (HEFLibs) for mutant rescue of p53 [15]. The halogen bond induced by the lead fragments proved to be crucial to the binding mode and affinity. Exchange of iodine by bromine and chlorine led to reductions in affinity that correlated well to the calculations presented here. This demonstrates the applicability of the presented theortical data to molecular design and fragment-based screening.

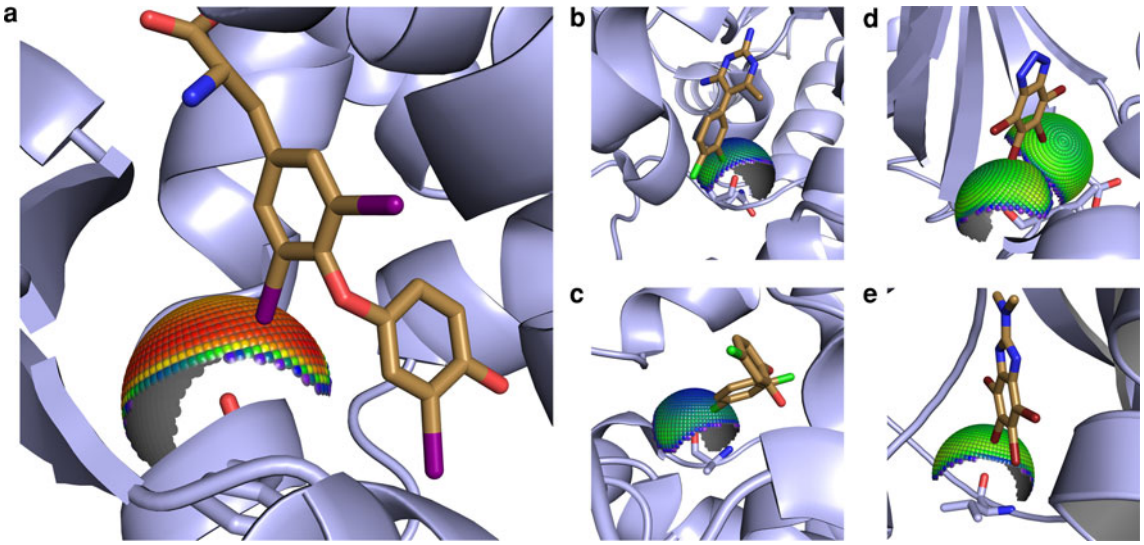


Fig. 8 Interaction spheres projected onto existing protein–ligand crystal structures. **a** pdb 1BSX (thyroid hormone receptor with 3.3′,5-triiodo-L-thyronine)[47]. **b** pdb 2AOV (histamine methyltransferase with metoprine)[48]. **c** pdb 2QIO (Bacterial enoyl-acyl carrier protein reductase with triclosan)[49]. **d** pdb 1P5E (phospho-CDK2/cyclin A with 4,5,6,7-tetrabromobenzotriazole)[50]. **e** pdb 1ZOE (protein kinase CK2 with TBB derivatives)[51]

Unfortunately, none of the commonly used molecular design tools recognizes halogen bonds as attractive contributions to binding energy at present. Based on systematic evaluations of halogen bonding based on this and similar studies, design algorithms can be developed and scoring functions can be trained.

A thorough understanding of molecular interactions is essential for medicinal chemistry and molecular design. With its potential and also its limitations, halogen bonding is a valuable asset to the drug discovery process and might promote halogens from being perceived as useful decorations to being part of the essential pharmacophore.

References

1. Hassel O (1970) Structural aspects of interatomic charge-transfer bonding. Science 170(3957):497–502. doi:10.1126/science.170.3957.497
2. Pierangelo M, Franck M, Tullio P, Giuseppe R, Giancarlo T (2008) Halogen bonding in supramolecular chemistry. Angew Chem Int Ed 47(33):6114–6127. doi:10.1002/anie.200800128
3. Auffinger P, Hays FA, Westhof E, Ho PS (2004) Halogen bonds in biological molecules. Proc Natl Acad Sci USA 101(48):16789–16794. doi:10.1073/pnas.0407607101
4. Clark T, Hennemann M, Murray J, Politzer P (2007) Halogen bonding: the σ-hole. J Mol Mod 13(2):291–296. doi:10.1007/s00894-006-0130-2
5. Murray J, Lane P, Politzer P (2009) Expansion of the σ-hole concept. J Mol Mod 15(6):723–729. doi:10.1007/s00894-008-0386-9
6. Murray JS, Lane P, Politzer P (2007) A predicted new type of directional noncovalent interaction. Int J Quantum Chem 107(12):2286–2292. doi:10.1002/qua.21352
7. Murray JS, Riley KE, Politzer P, Clark T (2010) Directional weak intermolecular interactions: σ-hole bonding. Aust J Chem 63(12):1598–1607. doi:10.1071/CH10259
8. Politzer P, Lane P, Concha M, Ma Y, Murray J (2007) An overview of halogen bonding. J Mol Mod 13(2):305–311. doi:10.1007/s00894-006-0154-7
9. Lu Y, Wang Y, Zhu W (2010) Nonbonding interactions of organic halogens in biological systems: implications for drug discovery and biomolecular design. Phys Chem Chem Phys 12(18):4543–4551. doi:10.1039/B926326H
10. Huber K, Brault L, Fedorov O, Gasser C, Filippakopoulos P, Bullock AN, Fabbro D, Trappe J, Schwaller J, Knapp S, Bracher F (2012) 7,8-dichloro-1-oxo-beta-carbolines as a versatile scaffold for the development of potent and selective kinase inhibitors with unusual binding modes. J Med Chem 55(1):403–413. doi:10.1021/jm201286z
11. Fedorov O, Huber K, Eisenreich A, Filippakopoulos P, King O, Bullock AN, Szklarczyk D, Jensen LJ, Fabbro D, Trappe J, Rauch U, Bracher F, Knapp S (2011) Specific CLK inhibitors from a novel chemotype for regulation of alternative splicing. Chem Biol 18(1):67–76. doi:10.1016/j.chembiol.2010.11.009
12. Hardegger LA, Kuhn B, Spinnler B, Anselm L, Ecabert R, Stihle M, Gsell B, Thoma R, Diez J, Benz J, Plancher J-M, Hartmann G, Banner DW, Haap W, Diederich F (2011) Systematic investigation of halogen bonding in protein–ligand interactions. Angew Chem Int Ed 50(1):314–318. doi:10.1002/anie.201006781
13. Xu Z, Liu Z, Chen T, Wang Z, Tian G, Shi J, Wang X, Lu Y, Yan X, Wang G, Jiang H, Chen K, Wang S, Xu Y, Shen J, Zhu W (2011) Utilization of halogen bond in lead optimization: a case study of rational design of potent phosphodiesterase type 5 (PDE5) inhibitors. J Med Chem 54(15):5607–5611. doi:10.1021/jm200644r
14. Hardegger LA, Kuhn B, Spinnler B, Anselm L, Ecabert R, Stihle M, Gsell B, Thoma R, Diez J, Benz J, Plancher JM, Hartmann G, Isshiki Y, Morikami K, Shimma N, Haap W, Banner DW, Diederich F (2011) Halogen bonding at the active sites of human cathepsin L and MEK1 kinase: efficient interactions in different environments. ChemMedChem 6(11):2048–2054. doi:10.1002/cmdc.201100353

15. Wilcken R, Liu X, Zimmermann MO, Rutherford TJ, Fersht AR, Joerger AC, Boeckler FM (2012) Halogen-enriched fragment libraries as leads for drug rescue of mutant p53. J Am Chem Soc 134(15):6810–6818. doi:10.1021/ja301056a
16. Joerger AC, Ang HC, Fersht AR (2006) Structural basis for understanding oncogenic p53 mutations and designing rescue drugs. Proc Natl Acad Sci USA 103(41):15056–15061. doi:10.1073/pnas.0607286103
17. Boeckler FM, Joerger AC, Jaggi G, Rutherford TJ, Veprintsev DB, Fersht AR (2008) Targeted rescue of a destabilized mutant of p53 by an in silico screened drug. Proc Natl Acad Sci USA 105(30):10360–10365. doi:10.1073/pnas.0805326105
18. Riley KE, Hobza P (2008) Investigations into the nature of halogen bonding including symmetry adapted perturbation theory analyses. J Chem Theory Comput 4(2):232–242. doi:10.1021/ct700216w
19. Riley KE, Murray JS, Politzer P, Concha MC, Hobza P (2009) Br–O complexes as probes of factors affecting halogen bonding: interactions of bromobenzenes and bromopyrimidines with acetone. J Chem Theory Comput 5(1):155–163. doi:10.1021/ct8004134
20. Wilcken R, Zimmermann MO, Lange A, Zahn S, Kirchner B, Boeckler FM (2011) Addressing methionine in molecular design through directed sulfur-halogen bonds. J Chem Theory Comput 7(7):2307–2315. doi:10.1021/Ct200245e
21. Lu Y, Shi T, Wang Y, Yang H, Yan X, Luo X, Jiang H, Zhu W (2009) Halogen bonding, a novel interaction for rational drug design? J Med Chem 52(9):2854–2862. doi:10.1021/jm9000133
22. Politzer P, Murray JS, Clark T (2010) Halogen bonding: an electrostatically-driven highly directional noncovalent interaction. Phys Chem Chem Phys 12(28):7748–7757. doi:10.1039/C004189K
23. Politzer P, Murray JS, Lane P (2007) Sigma-hole bonding and hydrogen bonding: competitive interactions. Int J Quantum Chem 107(15):3046–3052. doi:10.1002/qua.21419
24. Shields ZP, Murray JS, Politzer P (2010) Directional tendencies of halogen and hydrogen bonds. Int J Quantum Chem 110(15):2823–2832. doi:10.1002/qua.22787
25. Voth AR, Khuu P, Oishi K, Ho PS (2009) Halogen bonds as orthogonal molecular interactions to hydrogen bonds. Nat Chem 1(1):74–79. doi:10.1038/nchem.112
26. Weigend F, Ahlrichs R (2005) Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: design and assessment of accuracy. Phys Chem Chem Phys 7(18):3297–3305. doi:10.1039/B508541A
27. Ahlrichs R, Bär M, Häser M, Horn H, Kölmel C (1989) Electronic structure calculations on workstation computers: the program system turbomole. Chem Phys Lett 162(3):165–169. doi:10.1016/0009-2614(89)85118-8
28. TURBOMOLE v6.2 (2010), available from http://www.turbomole.com
29. Peterson KA, Figgen D, Goll E, Stoll H, Dolg M (2003) Systematically convergent basis sets with relativistic pseudopotentials. II. Small-core pseudopotentials and correlation consistent basis sets for the post-d group 16–18 elements. J Chem Phys 119(21):11113–11123. doi:10.1063/1.1622924
30. Feyereisen M, Fitzgerald G, Komornicki A (1993) Use of approximate integrals in ab initio theory. An application in MP2 energy calculations. Chem Phys Lett 208(5–6):359–363. doi:10.1016/0009-2614(93)87156-w
31. Weigend F, Häser M, Patzelt H, Ahlrichs R (1998) RI-MP2: optimized auxiliary basis sets and demonstration of efficiency. Chem Phys Lett 294(1–3):143–152. doi:10.1016/s0009-2614(98)00862-8
32. Hättig C (2005) Optimization of auxiliary basis sets for RI-MP2 and RI-CC2 calculations: core-valence and quintuple-z basis sets for H to Ar and QZVPP basis sets for Li to Kr. Phys Chem Chem Phys 7(1):59–66. doi:10.1039/B415208E
33. Hellweg A, Hättig C, Höfener S, Klopper W (2007) Optimized accurate auxiliary basis sets for RI-MP2 and RI-CC2 calculations for the atoms Rb to Rn. Theor Chem Acc 117(4):587–597. doi:10.1007/s00214-007-0250-5
34. Boys SF, Bernardi F (1970) The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. Mol Phys 19(4):553–566. doi:10.1080/00268977000101561
35. Grimme S (2003) Improved second-order Møller-Plesset perturbation theory by separate scaling of parallel- and antiparallel-spin pair correlation energies. J Chem Phys 118(20):9095–9102. doi:10.1063/1.1569242
36. Dunning JTH (1989) Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. J Chem Phys 90(2):1007–1023. doi:10.1063/1.456153
37. Woon DE, Dunning JTH (1993) Gaussian basis sets for use in correlated molecular calculations. III. The atoms aluminum through argon. J Chem Phys 98(2):1358–1371. doi:10.1063/1.464303
38. Peterson KA, Shepler BC, Figgen D, Stoll H (2006) On the spectroscopic and thermochemical properties of ClO, BrO, IO, and their anions. J Phys Chem A 110(51):13877–13883. doi:10.1021/jp065887l
39. Weigend F, Kohn A, Hättig C (2002) Efficient use of the correlation consistent basis sets in resolution of the identity MP2 calculations. J Chem Phys 116(8):3175–3183. doi:10.1063/1.1445115
40. Werner H-J, Knowles PJ, Lindh R, Manby FR, Schütz M, Celani P, Korona T, Rauhut G, Amos RD, Bernhardsson A, Berning A, Cooper DL, Deegan MJO, Dobbyn AJ, Eckert F, Hampel C, Hetzer G, Lloyd AW, McNicholas SJ, Meyer W, Mura ME, Nicklass A, Palmieri P, Pitzer R, Schumann U, Stoll H, Stone AJ, Tarroni R, Thorsteinsson T (2006) MOLPRO, version 2006.1, a package of ab initio programs; http://www.molpro.net
41. Hobza P, Šponer J (2002) Toward true DNA base-stacking energies: MP2, CCSD(T), and complete basis set calculations. J Am Chem Soc 124(39):11802–11808. doi:10.1021/ja026759n
42. Jurecka P, Sponer J, Cerny J, Hobza P (2006) Benchmark database of accurate (MP2 and CCSD(T) complete basis set limit) interaction energies of small model complexes, DNA base pairs, and amino acid pairs. Phys Chem Chem Phys 8(17):1985–1993. doi:10.1039/B600027D
43. Halkier A, Helgaker T, Jørgensen P, Klopper W, Koch H, Olsen J, Wilson AK (1998) Basis-set convergence in correlated calculations on Ne, N2, and H2O. Chem Phys Lett 286(3–4):243–252. doi:10.1016/s0009-2614(98)00111-0
44. DeLano WL (2008) The PyMOL molecular graphics system. DeLano Scientific LLC, Palo Alto, CA
45. Riley KE, Murray JS, Fanfrlik J, Rezac J, Sola RJ, Concha MC, Ramos FM, Politzer P (2012) Halogen bond tunability II: the varying roles of electrostatic and dispersion contributions to attraction in halogen bonds. J Mol Model. doi:10.1007/s00894-012-1428-x
46. Hennemann M, Murray JS, Politzer P, Riley KE, Clark T (2012) Polarization-induced sigma-holes and hydrogen bonding. J Mol Mod 18(6):2461–2469. doi:10.1007/s00894-011-1263-5
47. Darimont BD, Wagner RL, Apriletti JW, Stallcup MR, Kushner PJ, Baxter JD, Fletterick RJ, Yamamoto KR (1998) Structure and specificity of nuclear receptor–coactivator interactions. Genes Dev 12(21):3343–3356. doi:10.1101/gad.12.21.3343
48. Horton JR, Sawada K, Nishibori M, Cheng X (2005) Structural basis for inhibition of histamine N-methyltransferase by diverse drugs. J Mol Biol 353(2):334–344. doi:10.1016/j.jmb.2005.08.040
49. Tipparaju SK, Mulhearn DC, Klein GM, Chen Y, Tapadar S, Bishop MH, Yang S, Chen J, Ghassemi M, Santarsiero BD, Cook JL, Johlfs

M, Mesecar AD, Johnson ME, Kozikowski AP (2008) Design and synthesis of aryl ether inhibitors of the Bacillus Anthracis enoyl-ACP reductase. ChemMedChem 3(8):1250–1268. doi:10.1002/cmdc.200800047

50. De Moliner E, Brown NR, Johnson LN (2003) Alternative binding modes of an inhibitor to two different kinases. Eur J Biochem 270(15):3174–3181. doi:10.1046/j.1432-1033.2003.03697.x

51. Battistutta R, Mazzorana M, Sarno S, Kazimierczuk Z, Zanotti G, Pinna LA (2005) Inspecting the structure-activity relationship of protein kinase CK2 inhibitors derived from tetrabromo-benzimidazole. Chem Biol 12(11):1211–1219. doi:10.1016/j.chembiol.2005.08.015