J Comput Aided Mol Des (2015) 29:37–46
DOI 10.1007/s10822-014-9802-7

Molecular dynamics simulation of halogen bonding mimics experimental data for cathepsin L inhibition

Cristian Celis-Barros · Leslie Saavedra-Rivas · J. Cristian Salgado · Bruce K. Cassels · Gerald Zapata-Torres

Received: 26 March 2014 / Accepted: 10 October 2014 / Published online: 22 October 2014

Abstract A MD simulation protocol was developed to model halogen bonding in protein–ligand complexes by inclusion of a charged extra point to represent the anisotropic distribution of charge on the halogen atom. This protocol was then used to simulate the interactions of cathepsin L with a series of halogenated and non-halogenated inhibitors. Our results show that chloro, bromo and iodo derivatives have progressively narrower distributions of calculated geometries, which reflects the order of affinity I > Br > Cl, in agreement with the IC_{50} values. Graphs for the Cl, Br and I analogs show stable interactions between the halogen atom and the Gly61 carbonyl oxygen of the enzyme. The halogen-oxygen distance is close to or less than the sum of the van der Waals radii; the C–X···O angle is about 170°; and the X···O=C angle approaches 120°, as expected for halogen bond formation. In the case of the iodo-substituted analogs, these effects are enhanced by introduction of a fluorine atom on the inhibitors' halogen-bonding phenyl ring, indicating that the electron withdrawing group enlarges the σ-hole, resulting in improved halogen bonding properties.

Keywords Halogen bonding · MD simulation · Halogenated inhibitors · Cathepsin L · Interactions

Electronic supplementary material The online version of this article (doi:10.1007/s10822-014-9802-7) contains supplementary material, which is available to authorized users.

C. Celis-Barros · L. Saavedra-Rivas · J. C. Salgado · B. K. Cassels · G. Zapata-Torres (✉)
University of Chile, Santiago, Chile
e-mail: gzapata@uchile.cl

B. K. Cassels
e-mail: bcassels@uchile.cl

Introduction

As recently defined by the IUPAC, "A halogen bond occurs when there is evidence of a net attractive interaction between an electrophilic region associated with a halogen atom in a molecular entity and a nucleophilic region in another, or the same, molecular entity. A typical halogen bond is denoted by the three dots in R–X···Y" [1], where X is a halogen atom or a halonium ion, and Y is a halogen bond acceptor, typically an atom possessing a lone pair, an anion, or a π system. Such non-covalent interactions arise due to the anisotropic distribution of charge density on the halogen atom, resulting in a positively charged region at the far end of the R–X bond [2, 3]. Halogen bonds have energies comparable to those of hydrogen bonds (1–40 kcal/mol) [3], and their strength grows in the order Cl < Br < I. Fluorine is seldom able to form halogen bonds because it usually lacks a sufficiently positive region on the R–X bond axis, the so-called 'sigma hole' [4], although it can function as a halogen bond donor if it is bound to residues with strongly electron withdrawing groups [5, 6]. Halogen bonds, like hydrogen bonds, are strongly directional, with an optimal R–X···Y angle of about 180°, and the X···Y internuclear distance is typically less than the sum of the van der Waals radii of the participating atoms [4]. In addition, when Y is an atom bearing a lone electron pair, the X···Y–W angle should reasonably be close to the hybridization angle of atom Y. In the frequent case of a peptide carbonyl oxygen, this angle is expected to be close to 120°. Indeed, a theoretical study of the interactions between chloro-, bromo- and iodobenzene as halogen bond donors, and *N*-methylacetamide as the acceptor, gave values of 171.2°, 177.4° and 175.2° for the R–X···O angle, and 106.7°, 116.7° and 119.4° for the X···O=C angle, respectively. The same publication

calculated the X···O internuclear distances as 95.4, 90.2 and 87.4 % of the sum of the van der Waals radii of these halogens and oxygen [7]. Many theoretical and experimental studies have been done on halogen bonding, initially because of its possible applications in crystal engineering and in the development of solid electronic materials and polymers [8, 9]. Halogen bonds have also been observed fairly recently in biological systems [10], and they are beginning to be considered in rational drug design [11–14]. Nevertheless, this aspect can still be considered to be in its infancy despite the high likelihood of such interactions occurring in environments such as the carbonyl-rich protein active sites [7].

The electron density anisotropy involved in halogen bonds can be correctly described with quantum–mechanical methods such as DFT or MP2, but such methods are not applicable to large systems like halogenated drug molecules in biological systems. Computational docking protocols rely on force fields that incorporate a wide array of interactions, but do not include halogen bonds. The first paper to attempt modeling this interaction, introducing the concept of a point charge located on the X···Y axis, was published in 2011 [15] followed shortly thereafter by another one using a similar approach [16]. Other methodologies are being developed with the expectation that they will be able to provide conclusive evidence for the participation of these interactions in the affinity of specific halogenated molecules with their biological targets. Thus, several groups have developed modified force fields that have been tested in small models, though not yet with specific biological systems [17–22]. Endeavors of this sort naturally require biochemical-pharmacological data allowing experimental affinities to be compared with theoretically calculated parameters. Unfortunately, such data are for the most part scarce and incomplete. However, an exceptionally useful series of biological results surfaced in 2011, quantifying the biological activity of two extensive series of molecules that only differ in the presence of hydrogen, a methyl group, fluorine, chlorine, bromine or iodine at a specific location [12]. Some of these compounds, inhibitors of human cathepsin L, were also co-crystallized with the enzyme demonstrating most convincingly the formation of halogen bonds with a specific amino acid residue [12, 23].

Cathepsins are cysteine proteases that have a key role in physiological and pathological events. They are viewed as potentially useful drug targets, as their catalytic activity can be inhibited by small molecules that have become an attractive field for the pharmaceutical industry [24–27]. Cathepsin L is a lysosomal protease that is the main catalyst in intracellular protein catabolism, is associated with atherosclerosis, metabolic syndrome and cancer, and has also been related to bone resorption, sperm cell maturation,

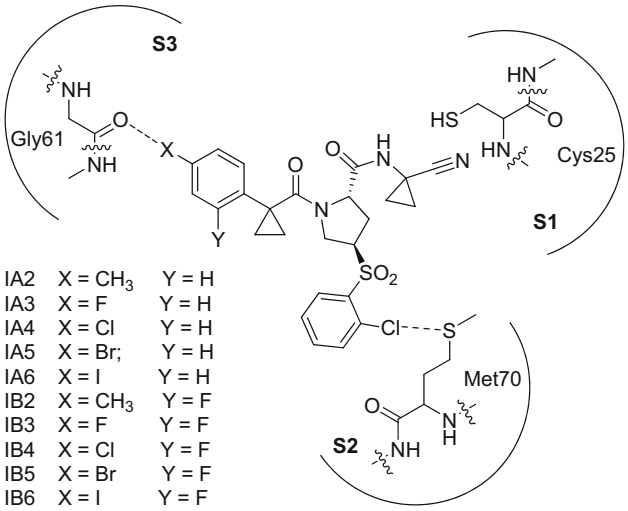

Fig. 1 Structures of the human cathepsin L inhibitors studied in this paper and their interactions with the enzyme, adapted from Ref. [12]. *Dotted lines* denote halogen bond interaction

macrophage malfunction and Alzheimer's disease [28–31]. A recently discovered family of cathepsin L inhibitors binds covalently to Cys25 in the S1 pocket of the enzyme and is further (or perhaps initially) stabilized in this location by hydrogen bonds involving Gly68 (in S2) and Asp162 (in S3) [32]. One of these inhibitors, bearing a chlorophenyl group in the region that binds in the S3 pocket, proved to be more than an order of magnitude more potent than its analog with an unsubstituted phenyl group. A collaboration between the Hoffmann-La Roche team and F. Diederich's group at the ETH led to the hypothesis that a halogen bond might be implicated in the chloro compound's enhanced affinity. The authors then proceeded to synthesize and test a wide array of derivatives of the parent structure and a related series, in all of which the phenyl (or pyridyl, or thienyl) ring expected to interact with the S3 pocket was unsubstituted or bore a methyl group, or a fluorine, chlorine, bromine or iodine atom and, in some instances, a second substituent (see Fig. 1).

Introduction of a chlorine atom at the *para* position of the phenyl ring or at C6 of a 3-pyridyl ring lowered the IC_{50} values (i.e., raised the affinities) several fold, and in a couple of instances by approximately 20 times. Moreover, in those cases in which the other abovementioned substituents were introduced, methyl and fluorine had little effect or even disfavored binding, while in three out of four series bromine was more favorable than chlorine and, in two of them, iodine was even more so than bromine. For both series of analogs the authors calculated that the gain in binding free energy on going from an unsubstituted phenyl to a *para*-chlorophenyl group was about 1.5 kcal/mol. The X-ray crystal structure of one of the *para*-chloro compounds showed the chlorine atom 3.1 Å away from the Gly61 backbone oxygen (the sum

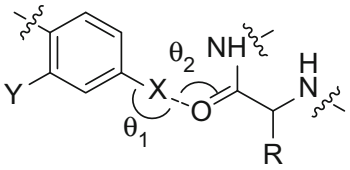

Fig. 2 Schematic representation of the characteristic angles θ_1 and θ_2 involved in a halogen bond between an aryl halide and an amino acid residue

of the van der Waals radii in this case is 3.27 Å), forming a C–Cl···O angle of 174°, lending further support to the hypothesis that a halogen bond participates in the stabilization of the enzyme-antagonist complex [12], and more X-ray results were added almost immediately, pointing in the same direction [23].

In view of the need to predict halogen bonds in the context of drug design, we built on the "extra point" or "explicit σ-hole" concept [15, 16], to model the halogen bond in drug-receptor or drug-enzyme interactions, and have applied it successfully to the human cathepsin L inhibitors depicted in Fig. 1. Using this methodology we managed to mimic the progressively shortened X···O distances (shorter than the sum of the van der Waals radii) between the inhibitor and a backbone oxygen atom when X=Cl, Br, or I, the near-optimal C–X···O angles θ_1, close to 180°, and also the X···O=C angles θ_2, which approach the expected value of 120° (Fig. 2). In addition, introduction of a fluorine atom at the *meta* position of the Cl-, Br-, or I-substituted ring enhances these effects as a consequence of the enlargement of the σ-hole. Taken together, all these features explain the increasing affinities of the different halogenated cathepsin L inhibitors on going from chlorine to the heavier halogen atoms [12].

Methods

In this study, a charged extra point was used to represent the anisotropic distribution of charge on the halogen atom [15]. Molecular dynamics calculations were then performed for cathepsin L inhibitors with the general structure shown in Fig. 1, where X represents H, CH_3, F, Cl, Br, or I and Y represents H or F.

Optimization of the halogen: Lewis base distance (quantum mechanical calculations)

In order to calculate the equilibrium distance between the implicated halogen atom of the inhibitors and the nucleophilic partner (the oxygen atom of Gly61 in this case), the complex consisting of a halobenzene (substituted by Cl, Br, I) and a formaldehyde molecule was chosen as a model system. The values obtained were used as references for molecular mechanics (MM) calculations carried out using Gaussian98. Due to the long distance reported between the Cl atom and the oxygen atom [15], it was necessary to use two different DFT functionals: the MPWLYP functional was employed for chlorine and B3LYP was used for Br and I. To obtain a better representation of the polarization for Br and I, the augmented correlation, consistent valence double-zeta basis set with a polarization function (aug-cc-pVDZ-PP), was used for these atoms. All the other atoms (C, H, O and Cl) were represented with the Pople triple-zeta basis set with polarization and diffuse functions (6-311+G**).

Inhibitor modeling

The chloro-substituted IA4 was extracted from the crystal structure of its complex with the enzyme (PDB code: 2xu1). In this particular case, the ligand is covalently bonded to the Cys25 residue. Therefore, it was necessary to consider a cysteine moiety bonded to the inhibitor in order to simulate this covalent bond. These considerations were taken into account in order to obtain the best partial charges and therefore better molecule parameterization.

Partial charge and RESP charge calculations

Once all the inhibitor molecules had been modelled, a ligand-cysteine complex optimization was carried out at the B3LYP/6-31G** level of theory excluding the halogen atom, which was treated using the aug-cc-pVDZ-PP basis set. In this manner the lowest energy geometry and the ESP charges were obtained. The ESP charges were extracted from Gaussian98 results using the ESPGEN program in AMBER. Due to the need to represent the σ hole, a massless dummy atom was introduced manually. This extra point was fixed at an optimized distance taken from Ref. [15] and then was subjected to the two-step RESP procedure implemented in ANTECHAMBER, which assigns a final partial positive charge that simulates the σ hole. Parameters for the inhibitors were calculated with the ANTECHAMBER program. In the case of the extra point, the parameters described in Table 1 of Ref. [15] were taken into account. The complex was then prepared for molecular dynamics runs, solvating the system with an octahedral box with a 10 Å radius protein cutoff. Finally, the anionic complex was neutralized with Na^+ atoms.

Molecular dynamics simulations

SANDER implemented in AMBER was the program used to carry out the MD simulations. The process was divided into four steps. The purpose of the first two was to relax the

Table 1 Sum of van der Waals radii (ΣvdW) [34]

Atom 1	Atom 2	ΣvdW (Å)
F	O	2.99
Cl	O	3.27
Br	O	3.35
I	O	3.50
C	O	2.15

system and search for local minimal conformational energies. Then, an equilibrating step was taken consisting of a gradual temperature increase, ending with a production step under equilibrium conditions.

1. The system was relaxed gradually in order to avoid unwanted atomic motions. All protein and ligand atoms were fixed by assigning a force constant (100 kcal/mol) that only left water molecules and sodium atoms free to move. This procedure was done over 250 initial steps of steepest descent and then 500 steps of conjugate gradient with a 10 Å cutoff.
2. Once water molecules and counterions had been located, the whole system was set free to move and then minimized over 500 steps.
3. In the equilibration step the system was heated from 0 to 300 K. The Langevin dynamics for temperature control program was used to increase the temperature gradually.
4. When the system had reached equilibrium, it was subjected to 10 ns full free NPT ensemble MD simulation.

Results and discussion

The cathepsin L inhibitors studied (Fig. 1) can be divided into two series. In the A series only the substituent at position X varies, allowing comparison of the effects of different halogen atoms and the bioisosteric methyl group on binding to the S3 pocket of the enzyme. The B series is additionally substituted with a *m*-fluorine atom at position Y, which may be expected to modify the strength of the halogen bond [33].

Observed inhibitor interactions

10 ns MD runs were carried out for all the inhibitors displayed in Table 1. The geometrical parameters analyzed were the distances from the variable substituent to the oxygen atom of Gly61, and the angles formed (a) by the aromatic carbon, its substituent, and the oxygen atom of Gly61, and (b) by the substituent, the Gly61 oxygen and the carbonyl carbon, as shown in Fig. 2.

A useful parameter to assess the existence of a specific interatomic interaction such as the halogen bond is the sum of the van der Waals radii (Table 1) of the participating atoms, which represents the minimum distance to which both atoms can approach without forming a bond. Thus, any distance shorter than the sum of the van der Waals radii of oxygen and a halogen may be interpreted as a strong indication of halogen bonding. For this purpose we have used the values from Bondi's compilation which, for the most part, have been confirmed by later studies [34]. According to the literature, the expected C–X···O angle should be close to 180°. Angles ranging from 160° to 180° are considered acceptable to decide if the interaction corresponds to a halogen bond. Similarly, the optimal X···O=C angle should be 120° considering the valence orbital hybridization of the oxygen atom.

In the case of the fluoro cathepsin L inhibitors IA3 and IB3, the histograms obtained from MD runs (Fig. 3 and SI1) show a range of F···O distances clustering around 3.2 Å, exceeding the sum of the van der Waals radii and suggesting that, as reported in the literature, there is no halogen bond interaction. These fluorine-substituted inhibitors show C–F···O angles fluctuating widely between 100° and 180°, as seen in Fig. 4 and SI2. The modes of the C–F···O angle distributions are about 152° and 158° for IA3 and IB3, respectively, smaller than the likely values for halogen bonds. It may be pointed out that the experimental IC_{50} values are almost identical (0.34 μM for IA3 and 0.35 μM for IB3, respectively), somewhat higher than for their counterparts lacking a halogen at the relevant position and even lower than the methyl analogs IA1 and IA2 [12], indicating the lack of any favorable interaction due to the fluorine atom.

The chloro analogs IA4 and IB4 show clearly narrower distributions of Cl···O distances (Fig. 3 and SI1) with modes around 3.2–3.3 Å, very close to the sum of the van der Waals radii. The plots of the C–Cl···O angles for these inhibitors (Figs. 4 and S2) show rather wide fluctuations, but with minima around 120°, and most populated values about 175° and 170° for IA3 and IB3, respectively. The latter is a clear indication of a fairly strong interaction that is reflected in the experimental IC_{50} values which are an order of magnitude lower than for the fluoro compounds [12].

In the case of the bromo inhibitors IA5 and IB5, the Br···O distances are even more narrowly distributed around 3.2 Å (Fig. 3 and SI1), about 95 % of the sum of the van der Waals radii, so now the implication of a halogen bond seems quite clear. Moreover, the C–Br···O angles show a smaller variation clustering around 176° (Figs. 4 and S2). Interestingly, comparison of the IC_{50} values of the chloro

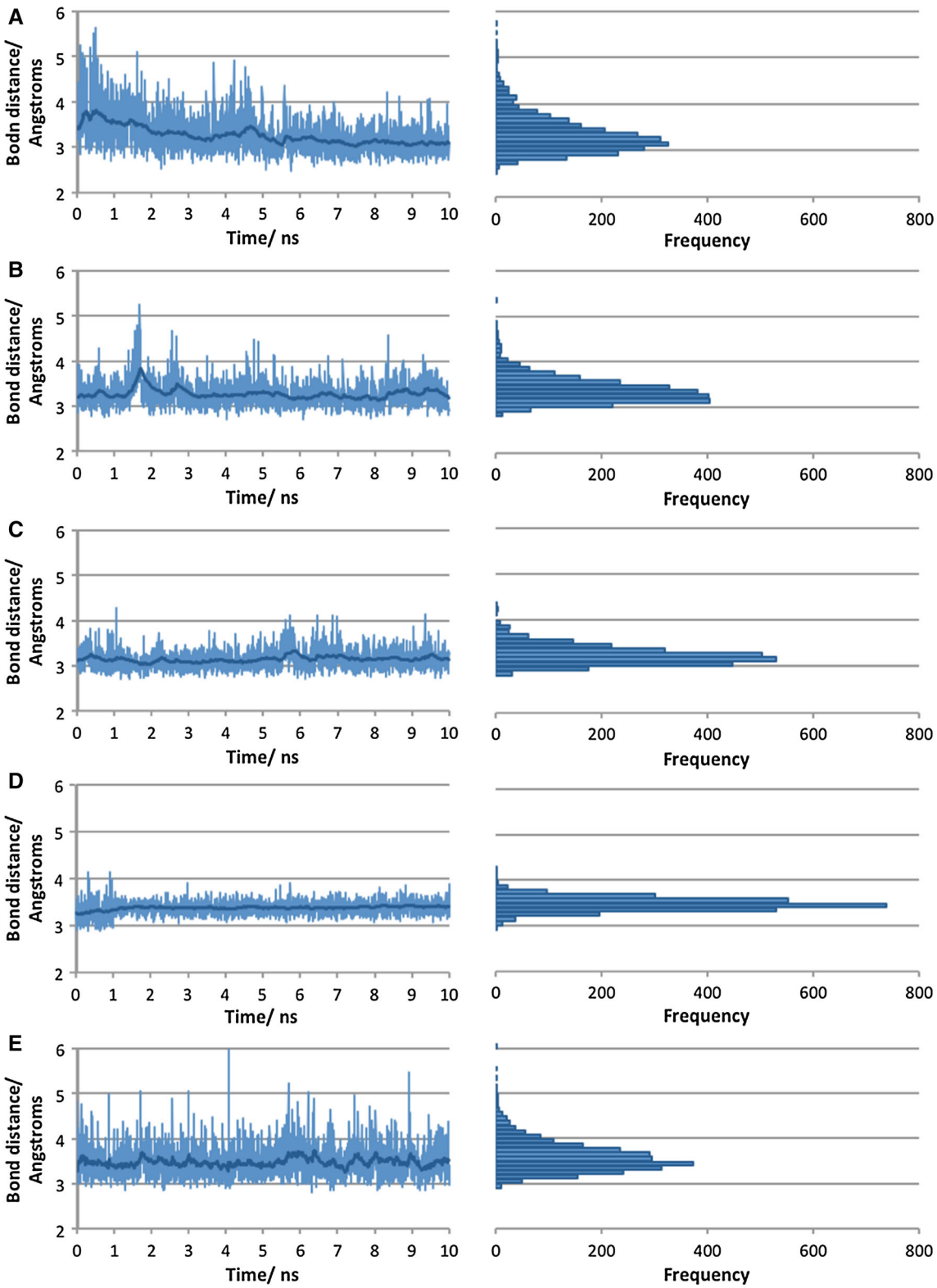

Fig. 3 MD trajectory plots (*left*) and histograms (*right*) for the X···O distances in cathepsin L-inhibitor complexes. **a** IA3 (F); **b** IA4 (Cl); **c** IA5 (Br); **d** IA6 (I); **e** IA2 (CH_3)

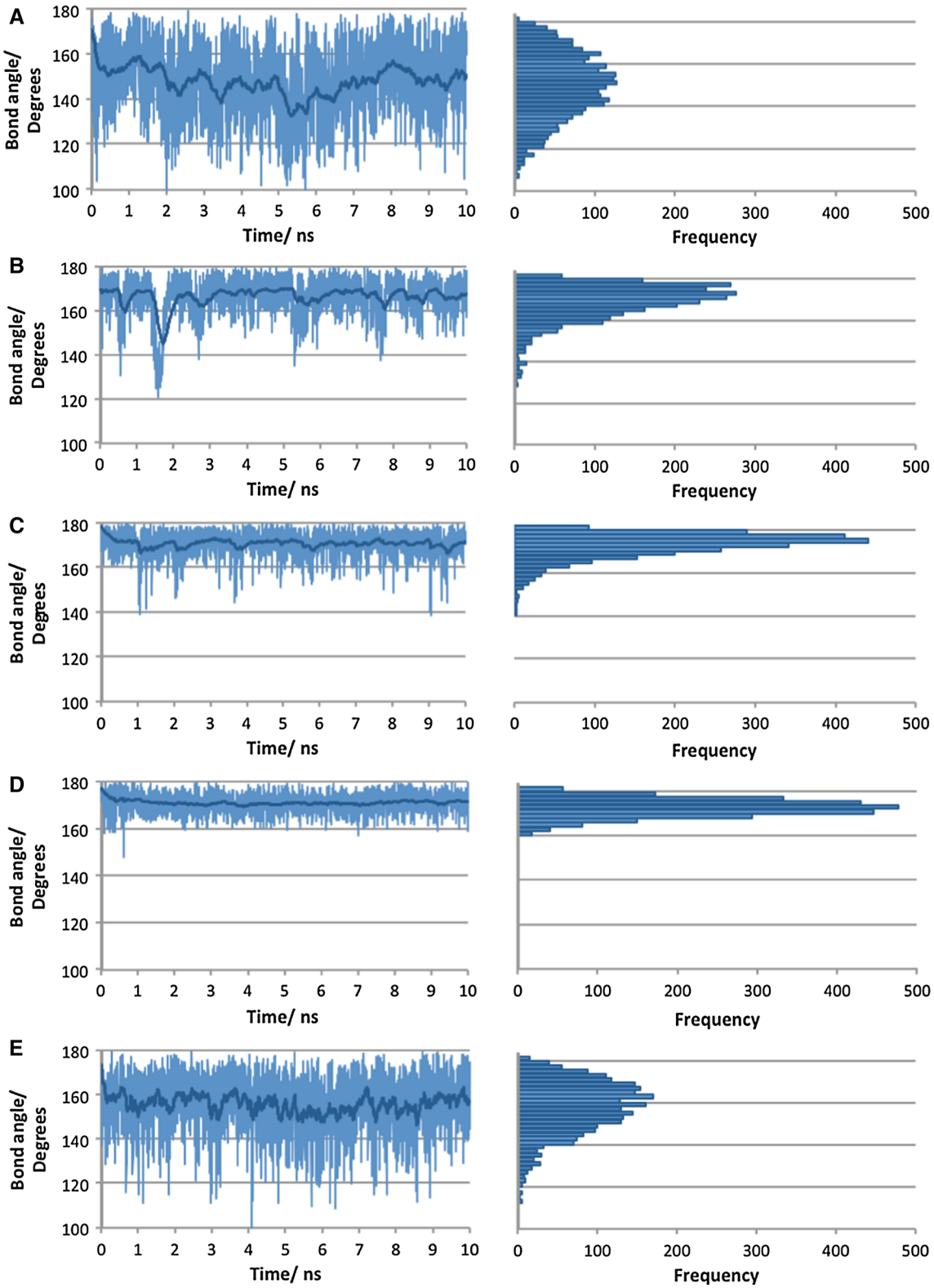

Fig. 4 MD trajectory plots (*left*) and histograms (*right*) for the C–X···O angles in cathepsin L-inhibitor complexes. **a** IA3 (F); **b** IA4 (Cl); **c** IA5 (Br); **d** IA6 (I); **e** IA2 (CH_3)

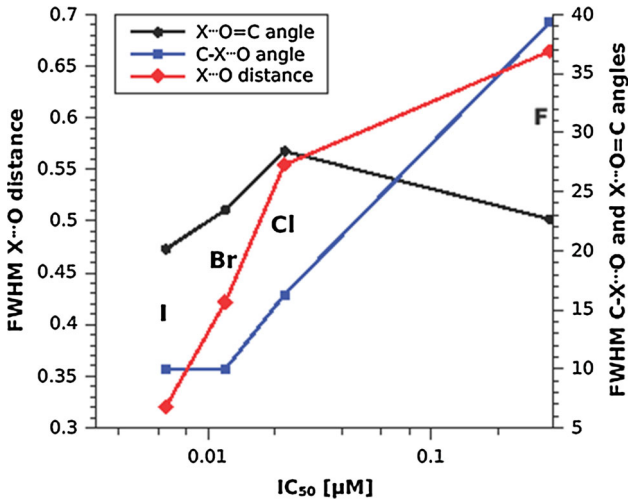

Fig. 5 Full width at half maximum (FWHM) values for the distributions of X···O distances, C–X···O and X···O=C angles observed during the simulation trajectories versus IC_{50} values (*logarithmic scale*) of compounds IA3 (F), IA4 (Cl), IA5 (Br) and IA6 (I)

and bromo compounds indicates that the potency of IA5 is approximately double that of IA4, while the introduction of a *m*-fluoro substituent is associated with an at least fourfold increase in potency on going from IB4 to IB5, suggesting that the strongly electronegative fluorine atom at the *meta* position might significantly favor a halogen bond involving bromine, but not chlorine.

The MD study of the iodo analogs continues and deepens the previous trend. Thus, the I···O distances are very narrowly distributed around 3.4 and 3.2 Å in the two series of compounds (Fig. 3 and SI1), 97 and 91 % of the sum of the van der Waals radii, and the C–I···O angles exhibit modes around 172° and 176° (Figs. 4 and S2). Here, visual comparison of the histograms for IA6 and IB6 reveals narrower I···O distance and C–I···O angle distributions of IB6 versus IA6, as well as their trend toward a shorter distance and a larger angle for the *m*-fluoro-substituted compound. The calculated full widths at half height of the corresponding histograms (Fig. 5, red trace for the X···O distances, blue trace for the C–X···O angles), clearly illustrate and quantify the trend to narrower distributions on going from F to Cl, to Br and I. The IC_{50} values for the iodinated compounds show higher potency than for their bromo counterparts, and the value for IB6, which is lower than for IA6, again supports the idea that *meta* fluorine substitution favors halogen bonding via a less electronegative, more polarizable atom like iodine [12].

Substitution with a methyl group, whose size is very similar to that of a chlorine atom, is a good test to determine if the interactions between the ligand and protein in the vicinity of Gly61 are related preferentially to the size of the S3 cavity in the cathepsin L molecule or to a specific interaction associated with the charge density anisotropy of the halogen atom acting as a putative Lewis acid. The IC_{50} value for the methylated IA2 (0.13 μM) is several times higher than for the *chloro*-substituted IA4 (0.022 μM). The MD simulations of the methyl-substituted inhibitors interacting with cathepsin L, either with or without additional *m*-fluoro substitution (Fig. 3 and SI1, Fig. 4 and SI2), display wider interatomic distance and angular distributions than for the complexes with the chloro-substituted ligands, although the modes of the C–Cl···O angles are slightly higher than 160°, so it seems clear that the higher potency of the chlorinated inhibitors with regard to IA2 (IB2 does not appear to have been synthesized or tested) is largely dictated by formation of a halogen bond [12].

As mentioned in the introduction, the X···O=C angle should be expected to be close to 120° if a halogen bond is implicated. Our results (SI3) are not as clear-cut as in the case of the C–X···O angles and internuclear distances, as the most populated angles lie between 132° and 142° with no obvious trend related to the identity of the halogen. Nevertheless a progressive narrowing of the angle distributions is quite obvious for the cases of the bromo and iodo compounds (IA5, IA6, IB5, and IB6), and again may be taken as indicative of an interaction that reduces the mobility of the inhibitor in the enzyme's S3 pocket. This is further illustrated in Fig. 5, black trace, showing the full widths at half height of the respective histograms.

The premise of the analyses in this study is that narrower distributions of geometries (distances and angles) are expected to be correlated with more enthalpically favorable interactions. Although a narrower distribution means the loss of translational and rotational degrees of freedom, it has been recently reported [35] that the TΔS term (from the Gibbs free energy equation) approaches a limit value and, in the case of stronger interactions, the ΔH value surpasses the entropy term resulting in a thermodynamically stable interaction.The reported IC_{50} values for cathepsin L might well be related to this stronger interaction. Figure 5 shows that, advancing from right to left in the -F, -Cl, -Br, -I sequence from higher IC_{50} values to lower ones (and therefore to higher potencies), the variation observed in the molecular dynamics simulation trajectories for the X···O distances and the C–X···O and X···O=C angles decreases. Thus, the smallest variations are seen in the case of iodine, and the largest ones for fluorine, with the exception of the possibly less exacting X···O=C angle, where F appears to be as good as Br. A smaller halogen atom is generally associated with greater mobility and higher entropy, but in the cases studied here the increasing enthalpic aspect of the halogen bond involving larger atoms more than compensates for this effect. As in this series fluorine is not expected to form a halogen bond, there is no balancing out

Table 2 50 % inhibitory concentrations of cathepsin L inhibitors and comparison of experimental [12, 23] and calculated geometrical parameters of ligand–protein interactions

Ligand	IC_{50} (μM)	X···O distance (Å)		C–X···O angle (°)		X···O=C angle (°)	
		X-ray	MD	X-ray	MD	X-ray	MD
IA2	0.13	3.6[a]	3.4	N.R.	160	N.R.	142
IA3	0.34	4.5[b]	3.5	N.R.	152	N.R.	138
IA4	0.022	3.0	3.2	171	172	147	136
IA5	0.012	3.1	3.2	176	175	146	132
IA6	0.0065	3.1	3.4	175	174	145	140

N.R. Not Reported

[a] C···O distance

[b] With an intervening water molecule

due to this interaction. It must be pointed out here that the co-crystal structure of cathepsin L with fluoro analog IA3 shows a water molecule bridging the fluorine atom and the Gly61 carbonyl oxygen, a feature that we did not include in our modeling studies.

Comparisons with crystallographic data

X-Ray structures are available for the cathepsin L complexes with several of the inhibitors studied here [12, 23]. Table 2 shows the experimental values for the relevant internuclear distances and angles together with the most populated values found by us in our MD runs. As can be seen in the table, the agreement between the experimental and calculated values is generally quite good. Even though the I···O distance in IA6 is appreciably greater than the crystallographic distance, it is still less than the sum of the van der Waals radii (Table 1). It should be pointed out here that optimal binding geometries are not generally observed in ligand–protein binding due to the interplay of competing interactions [7, 36]. Interestingly, the most populated MD values for the X···O=C angle are smaller, and therefore closer to the (in theory) ideal value of 120°, than those found in the crystal structures. This might be a reflection of the greater conformational mobility of the complexes in solution than in the crystal phase, as would be the case in vivo, in spite of the intrinsic limitations imposed by our model.

Secondary halogen bond

All the compounds examined in this study share an *o*-chlorobenzenesulfonyl moiety that binds in cathepsin L's S2 pocket, in close proximity to the Met70 sulfur atom [12]. This suggests the possibility of secondary halogen bond between the chlorine atom on the chlorobenzenesulfonyl group of the inhibitor (see Fig. 1) and the sulfur atom of a methionine residue in the enzyme S2 region. We therefore carried out a search the same as before, i.e., looking for a C–Cl···S angle close to 180° and an internuclear distance shorter than the sum of the van der Waals radii. SI5 and SI7 show that the C–Cl···S angles cluster around 160° or even less, which would seem to indicate that there is no halogen bond involved. The sum of the chlorine and sulfur van der Waals radii is 3.55 Å, and the most populated bond distances found for the methyl-, fluorine-, chlorine- and bromine-substituted inhibitors are greater than this value, in both series A and B (SI4 and SI6). Nevertheless, calculations using as models a small sulfur-containing molecule and chloro-, bromo- and iodobenzene suggest that C–Cl···S angles as small as 150° and Cl···S distances up to 3.89 Å may be tolerated [36]. The iodinated cathepsin L inhibitors IA6 and IB6, however, exhibit Cl···S distances that stabilize at ca. 3.50 Å or less, and the distribution of Cl···S internuclear distances and C–Cl···S angles in both cases is extremely narrow, another indication of a strong interaction that could reasonably be a halogen bond. Although the modes of the X···S distances and C–X···S angles do not lend clear support to the idea of halogen-methionine bonding of the chlorine- and bromine-substituted inhibitors in the S2 pocket, this hypothesis cannot be completely excluded as the MD runs and histograms for these ligands show visibly narrower distributions than for the methyl- or fluorine-substituted compounds, indicating reduced mobility in this area of the respective enzyme-inhibitor complexes. Indeed, it is rather surprising that a chlorine or iodine (but not bromine) atom at a site far removed from the Cl···S pair should have any noticeable effect on the stability of a distant and possibly weak halogen bond between a chlorine atom and the sulfur atom of Met70.

Conclusions

The methodology presented here, based on the introduction of an extra point charge to simulate the σ-hole predicts the halogen bonds observed in two series of cathepsin L inhibitors. It reproduces to a certain extent the geometrical parameters described for the series—Cl < Br < I—and the absence of such an interaction for inhibitors in which the atom facing the enzyme's S3 site is fluorine. For chlorine and bromine there is a reasonable agreement between the calculated halogen bond distances and angles with data obtained by X-ray diffraction analysis of the enzyme-inhibitor complexes. However, for the complexes with iodo compounds, the bond geometry is not close to the crystal structure. Our models also explain the favorable effect of a *meta*-fluorine atom on the binding affinity of ligands with more polarizable iodine atoms interacting with this site,

and suggest the existence of an additional secondary halogen bond in the S2 site.

As pointed out by Carter et al. [37], although halogen bonds involving iodine are enthalpically favored over those involving smaller halogen atoms, this situation may be more than compensated by an entropic penalty due to crowding effects in their DNA model, resulting in optimal stabilization by a Br···O instead of an I···O interaction. This does not seem to be our case where iodine, as shown experimentally [12], leads to the highest enzyme inhibitory potencies and, in the present study, to the (relatively) shortest X···O distances and most favorable C–X···O and X···O=C angles. Thus, the enthalpies from the halogen bonds in the complexes, as reflected in their geometries, are correlated to the IC_{50} values in these particular cathepsin L inhibitors. However, as stated above for other ligand systems, iodine substitutions may in fact be detrimental relative to bromine or chlorine. Overall, it is clear that our methodology should be useful to predict the occurrence of halogen bonding in other biological systems, thus enriching the drug designer's armamentarium in a novel way.

Acknowledgments This study was supported by FONDECYT Grant 1110146.

Conflict of interest The authors declare no competing financial interests.

References

1. Desiraju GR, Ho PS, Kloo L, Legon AC, Marquardt R, Metrangolo P, Politzer P, Resnati G, Rissanen K (2013) Definition of the halogen bond (IUPAC Recommendations 2013). Pure Appl Chem 85:1711–1713
2. Politzer P, Lane P, Concha MC, Ma Y, Murray JS (2007) An overview of halogen bonding. J Mol Model 13:305–311
3. Metrangolo P, Neukirch H, Pilati T, Resnati G (2005) Halogen bonding based recognition processes: a world parallel to hydrogen bonding. Acc Chem Res 38:386–395
4. Clark T, Hennemann M, Murray J, Politzer P (2007) Halogen bonding: the σ-hole. J Mol Model 13:291–296
5. Metrangolo P, Murray JS, Pilati T, Politzer P, Resnati G, Terraneo G (2011) The fluorine atom as a halogen bond donor, viz a positive site. Cryst Eng Comm 13:6593–6596
6. Metrangolo P, Murray JS, Pilati T, Politzer P, Resnati G, Terraneo G (2011) Fluorine-centered halogen bonding: a factor in recognition phenomena and reactivity. Cryst Growth Des 11:4238–4246
7. Wilcken R, Zimmermann MO, Lange A, Zahn S, Boeckler FM (2012) Using halogen bonds to address the protein backbone: a systematic evaluation. J Comput Aided Mol Des 26:935–945
8. Politzer P, Murray JS, Concha MC (2007) Halogen bonding and the design of new materials: organic bromides, chlorides and perhaps even fluorides as donors. J Mol Model 13:643–650
9. Metrangolo P, Resnati G (2008) Halogen bonding: fundamentals and applications. Springer, Berlin
10. Scholfield MR, Zanden CMV, Carter M, Ho PS (2013) Halogen bonding (X-bonding): a biological perspective. Protein Sci 22:139–152
11. 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 52:2854–2862
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 50:314–318
13. Xu Z, Liu Z, Chen T, 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 54:5607–5611
14. Wilcken R, Zimmermann MO, Lange A, Joerger AC, Boeckler FM (2013) principles and applications of halogen bonding in medicinal chemistry and chemical biology. J Med Chem 56:1363–1388
15. Ibrahim MAA (2011) Molecular mechanical study of halogen bonding in drug discovery. J Comput Chem 32:2564–2574
16. Rendine S, Pieraccini S, Forni A, Sironi M (2011) Halogen bonding in ligand-receptor systems in the framework of classical force fields. Phys Chem Chem Phys 13:19508–19516
17. Kolář M, Hobza P (2012) On extension of the current biomolecular empirical force field for the description of halogen bonds. J Chem Theory Comput 8:1325–1333
18. Jorgensen WL, Schyman P (2012) Treatment of halogen bonding in the OPLS-AA force field: application to potent anti-HIV agents. J Chem Theory Comput 8:3895–3901
19. Carter M, Rappé AK, Ho PS (2012) Scalable anisotropic shape and electrostatic models for biological bromine halogen bonds. J Chem Theory Comput 8:2461–2473
20. El Hage K, Piquemal J-P, Hobaika Z, Maroun RG, Gresh N (2013) Could an anisotropic molecular mechanics/dynamics potential account for sigma hole effects in the complexes of halogenated compounds? J Comput Chem 34:1125–1135
21. Du L, Gao J, Bi F, Wang L, Liu C (2013) A polarizable ellipsoidal force field for halogen bonds. J Comput Chem 34:2032–2040
22. Liu Y, Xu Z, Yang Z, Chen K, Zhu W (2013) A knowledge-based halogen bonding scoring function for predicting protein-ligand interactions. J Mol Model 19:5015–5030
23. 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, 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 6:2048–2054
24. Debouck C, Metcalf B (2000) The impact of genomics on drug discovery. Annu Rev Pharmacol Toxicol 40:193–207
25. Podgorski I (2009) Future of anticathepsin K drugs: dual therapy for skeletal disease and atherosclerosis? Future Med Chem 1:21–34
26. Huryn DM, Smith AB 3rd (2009) The identification, characterization and optimization of small molecule probes of cysteine proteases: experiences of the Penn Center for molecular discovery with cathepsin B and cathepsin L. Curr Top Med Chem 9:1206–1216
27. Schröder J, Klinger L, Oellien F, Marhöfer RJ, Duszenko M, Selzer PM (2013) Docking-based virtual screening of covalently-binding ligands: an orthogonal lead discovery approach. J Med Chem 56:1478–1490
28. Leto G, Sepporta MV, Crescimanno M, Flandina C, Tumminello FM (2010) Cathepsin L in metastatic bone disease: therapeutic implications. Biol Chem 391:655–664
29. Lafarge JC, Naour N, Clément K, Guerre-Millo M (2010) Cathepsins and cystatin C in atherosclerosis and obesity. Biochimie 92:1580–1586

30. Reiser J, Adair B, Reinheckel T (2010) Specialized roles for cysteine cathepsins in health and disease. J Clin Invest 120:3421–3431
31. Hook V, Funkelstein L, Wegrzyn J, Bark S, Kindy M, Hook G (2012) Cysteine cathepsins in the secretory vesicle produce active peptides: cathepsin L generates peptide neurotransmitters and cathepsin B produces beta-amyloid of Alzheimer's disease. Biochim Biophys Acta 1824:89–104
32. Álvarez Sánchez R, Banner D, Ceccarelli SM, Grether U, Haap W, Hartman P, Hartmann G, Hilpert H, Kuehne H, Mauser H, Plancher J-M. (F. Hoffmann-La Roche A.G.), Novel Proline Derivatives. US20100267722 (A1), October 21, 2010
33. Riley KE, Murray JS, Fanfrlík J, Řezáč J, Solá RJ, Concha MC, Ramos FM, Politzer P (2011) Halogen bond tunability I: the effects of aromatic fluorine substitution on the strengths of halogen-bonding interactions involving chlorine, bromine, and iodine. J Mol Model 17:3309–3318
34. Bondi A (1964) van der Waals volumes and radii. J Phys Chem 68:441–451
35. Politzer P, Murray JS (2013) Enthalpy and entropy factors in gas phase halogen bonding: compensation and competition. Cryst Eng Comm 15:3145–3150
36. Wilcken R, Zimmerman MO, Lange A, Zahn S, Kirchner B, Boeckler FM (2011) Addressing methionine in molecular design through directed sulphur-halogen bonds. J Chem Theory Comput 7:2307–2315
37. Carter M, Voth AR, Scholfield MR, Rummel B, Sowers LC, Ho PS (2013) Enthalpy–entropy compensation in biomolecular halogen bonds measured in DNA junctions. Biochemistry 52:4891–4903