

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Halogen Bonding in the framework of classical force fields: the case of chlorine

Davide Franchini,[a] Federico Dapiaggi,[a] Stefano Pieraccini,[a, b] Alessandra Forni,*[b] and Maurizio Sironi*[a,b]*

[a] Dr. D. Franchini, Dr. F. Dapiaggi, Dr. S. Pieraccini, Prof. Dr. M. Sironi

Dipartimento di Chimica and INSTM UdR, Università degli Studi di Milano

Via Golgi 19, 20133, Milano (Italy)

E-mail: stefano.pieraccini@unimi.it, maurizio.sironi@unimi.it

[b] Dr. A. Forni, Dr. S. Pieraccini, Prof. Dr. M. Sironi

CNR-ISTM (Istituto di Scienze e Tecnologie Molecolari del CNR) and INSTM UdR

Via Golgi 19, 20133, Milano (Italy)

E-mail: alessandra.forni@istm.cnr.it

ABSTRACT

Halogen bonding is nowadays a consolidated tool in chemistry. Only recently, the importance of halogen bonding has been demonstrated also in biological systems, owing to the presence of halogens in drugs. This interaction is due to the anisotropy of the electron density around the halogen that leads to the formation of the ‘ σ -hole’, which is responsible for the interaction with a nucleophile site. Unfortunately, classical force fields used in the study of ligand-receptor systems

1
2
3
4 are not able to describe the ‘ σ -hole’. Here, we propose a pseudo-atom based methodology able to
5
6 correctly describe halogen bonding involving chlorine using classical force field.
7
8
9

10 11 **INTRODUCTION**

12
13
14 Halogen bonding (XB) is a highly directional, non-covalent interaction established between a
15
16 halogen atom (X) covalently bonded to a donor group (D) and a nucleophilic site A, the halogen
17
18 bonding acceptor, according to the scheme D–X \cdots A. The donor moiety D can range from another
19
20 halogen atom to an organic or inorganic group, whereas the nucleophilic acceptor site is most
21
22 often a lone pair on a heteroatom such as nitrogen, oxygen or sulfur. At first glance, the
23
24 formation of halogen bonding appears a rather unexpected phenomenon as it involves the
25
26 interaction between an electronegative, supposed spherical, species and a Lewis base. Actually,
27
28 the electron density around a covalently bonded halogen atom is not uniformly distributed,
29
30 exhibiting a depletion along the D–X bond axis in the region outward the bond itself, that is
31
32 compensated by a relative increase in a belt around the bond axis. A simple and intuitive model
33
34 to explain the origin of such anisotropy has been developed by Politzer and coworkers.¹
35
36 Compared to an isolated halogen atom, which has spherical symmetry and 5/3 average p electron
37
38 population in each direction, the electronic structure of a halogen atom covalently bonded to
39
40 another atom along the z axis can be approximately written as $s^2p_x^2p_y^2p_z^1$. The region of lower
41
42 electron density along the extension of the D–X bond, denoted as σ -hole,² generally reflects in a
43
44 positive electrostatic potential which is responsible for the interaction with a nucleophile site.
45
46 The size and depth of the σ -hole depend on the electron withdrawing capability of the donor
47
48 group and on the nature of the halogen atom. A good correlation has been observed between the
49
50 values of the electrostatic potential (ESP) on the σ -hole and the strength of halogen bond,³
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 although the electrostatic contribution is not the only energetic term giving rise to this
5
6 interaction. Depending on the nature of the halogen and of the donor and acceptor groups, also
7
8 dispersive forces can play a significant role.^{4,5}
9

10
11 The persistent view of halogens as electronegative species has long prevented researchers to
12
13 think to bonded halogens as responsible for even strong interactions with nucleophilic sites. Only
14
15 in the past two decades, differently for the more explored hydrogen bonding (HB), the potential
16
17 of XB in directing assembly phenomena has become to be explored. Thanks to its strength,
18
19 directionality and intrinsic tunability, the latter resulting from the possibility to span the four
20
21 halogen atoms, halogen bonding has found a plethora of possible applications in the most
22
23 diversified fields of chemistry and biology, going from photoluminescent materials⁶ to nonlinear
24
25 optics,⁷ supramolecular chemistry^{8,9} and drug design.¹⁰⁻¹⁴ In addition, XB is ubiquitously present
26
27 in biological systems, as pointed out by protein data bank (PDB) analysis.¹⁸ For a long time,
28
29 halogenated drug molecules were predominantly adopted due to their increased metabolic
30
31 stability and membrane permeability conferred by halogen atoms. It has been also shown that
32
33 XB can compete with HB in stabilizing brominated DNA junctions.^{10,15} In recent years it has
34
35 been observed that halogen atoms can increase the drug-target affinity through XB, thus
36
37 increasing drug activity of several orders of magnitude with respect to analogous non-
38
39 halogenated moieties.^{11,16,17} For these reasons, halogen bonding has been the subject of
40
41 numerous computational studies on both molecular and macromolecular systems. At molecular
42
43 level, highly correlated methods and density functional theory (DFT) have been largely used to
44
45 study this interaction in gas-phase¹⁹⁻²² and in solution using continuum models of the solvent.²³⁻²⁵
46
47 Quantum mechanics/molecular mechanics (QM/MM) methods have been used to describe the
48
49 interaction of halogenated molecules with proteins.¹⁸ In this context, many efforts have been
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 made in the direction of correctly describe XB in the framework of classical force fields methods
5
6 that are world-wide used to perform drug discovery and drug design studies. The main problem
7
8 in classical force field description of halogen bonding is that computation of the electrostatic
9
10 contribution to the interaction (which is generally the dominating term in XB) is based on a set
11
12 of atomic-point charges, usually obtained exploiting a fitting procedure (Restrained ElectroStatic
13
14 Potential, RESP²⁶) to quantum mechanical ESP. This approach is completely misleading in the
15
16 description of halogen bonding, which is based on the anisotropy of the electrostatic potential
17
18 around the halogen atom (see the σ -hole concept illustrated above). Several solutions have been
19
20 proposed through the years to fix this issue.²⁷ Among these it is to be mentioned the introduction
21
22 of multipoles replacing atomic centered point charges (at least on the atoms involved in the
23
24 interaction), allowing a rigorous description of XB in biological systems at the cost of time
25
26 demanding simulations and more complicate model with respect to the point charges one²⁸ and
27
28 several force fields, like OPLS3^{29,30} and the extended versions of AMOEBA³¹ and CGenFF³²,
29
30 designed specifically to treat this kind of non-covalent interaction. Other possible solutions are
31
32 force fields allowing for aspherical halogen atoms³³⁻³⁵ or force fields where an extra-point charge
33
34 is introduced, directly linked to the halogen atom in the region outward the D–X bond, for the
35
36 purpose of simulating the charge induced by the presence of the σ -hole. This strategy has been
37
38 successfully implemented by different groups.³⁶⁻³⁸ In particular, Sironi and coworkers³⁸ applied it
39
40 to the study of brominated and iodinated ligands complexing proteins. Interestingly, the same
41
42 approach applied for the description of weaker halogen bonds involving chlorine resulted in
43
44 rather poor results (it has not been possible, differently from iodinated and brominated
45
46 compounds to reproduce crystallographic data involving chlorinated ligands). Indeed, chlorine is
47
48 the less polarizable among the halogens³⁹ (excluding fluorine that in most cases is not able to
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 give rise to halogen bonding⁴⁰) and therefore it is envisaged that it requires a more sophisticated
5
6 description of the ESP anisotropy around it. The necessity to improve the 1Pa methodology has
7
8 been also pointed out in other papers from MacKerrel et al.^{41, 42}
9

10
11 In this work, we propose a simple and effective way, based on the use of a set of extra-point
12
13 charges or pseudo-atoms (Pa's), properly placed around the halogen atom, to describe halogen
14
15 bonding involving chlorine in biological systems. The proposed approach has been applied to the
16
17 investigation of XB between a series of chlorinated model systems and the carbonyl oxygen of
18
19 the capped alanine (see Figure 1) through molecular mechanics (MM) geometry optimization.
20
21 One brominated derivative has been as well included to check the generality of our procedure.
22
23 These model systems have been chosen to simulate real halogenated ligands such as those
24
25 present in biological systems, as it has been shown that almost always, in the PDB, halogen
26
27 atoms involved in halogen bonding are linked to an aromatic ring.^{3,43} Aromatic, rather than
28
29 aliphatic, halogenated ligands are also important from an applicative point of view owing to the
30
31 greater strength of the associated C–X bonds, reducing the possible release of halogen anions in
32
33 the organism. The substituents on the aromatic ring, their position and the nature of the ring itself
34
35 have been varied to tune the XB strength in order to validate our methodology for a wide range
36
37 of halogen bonding interactions. We found out that the proposed approach allows to significantly
38
39 improve the geometrical parameters (i.e., interaction distances and angles) of the chlorine XB
40
41 interaction with respect to the previous pseudo-atom strategy,³⁶⁻³⁸ using as benchmark the results
42
43 obtained at M06-2X⁴⁴/6-311++G(d,p) level. Moreover, we show that a re-parametrization of the
44
45 general AMBER force field (GAFF⁴⁵) van der Waals parameters for chlorine is necessary to
46
47 have an effective description of chlorine XB in biological systems, in good agreement with the
48
49 results obtained by Scholfield et al.³⁴
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

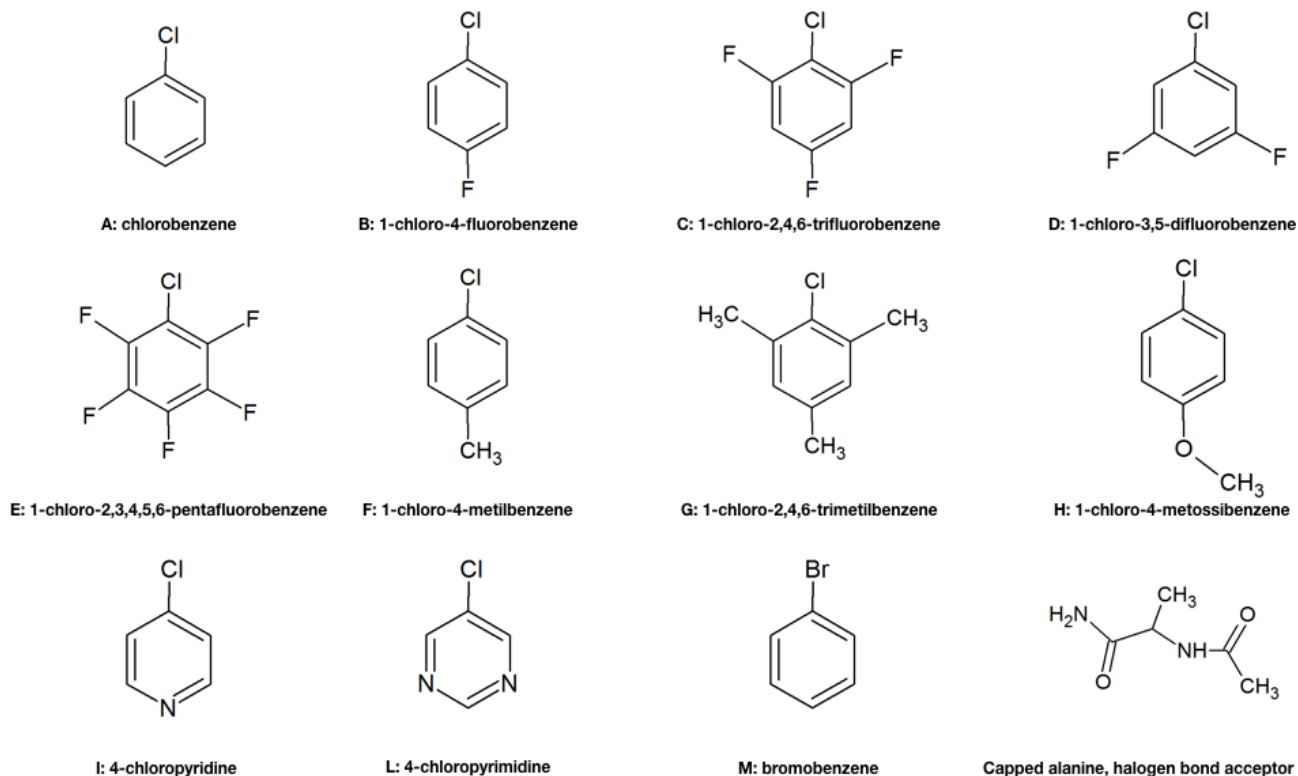


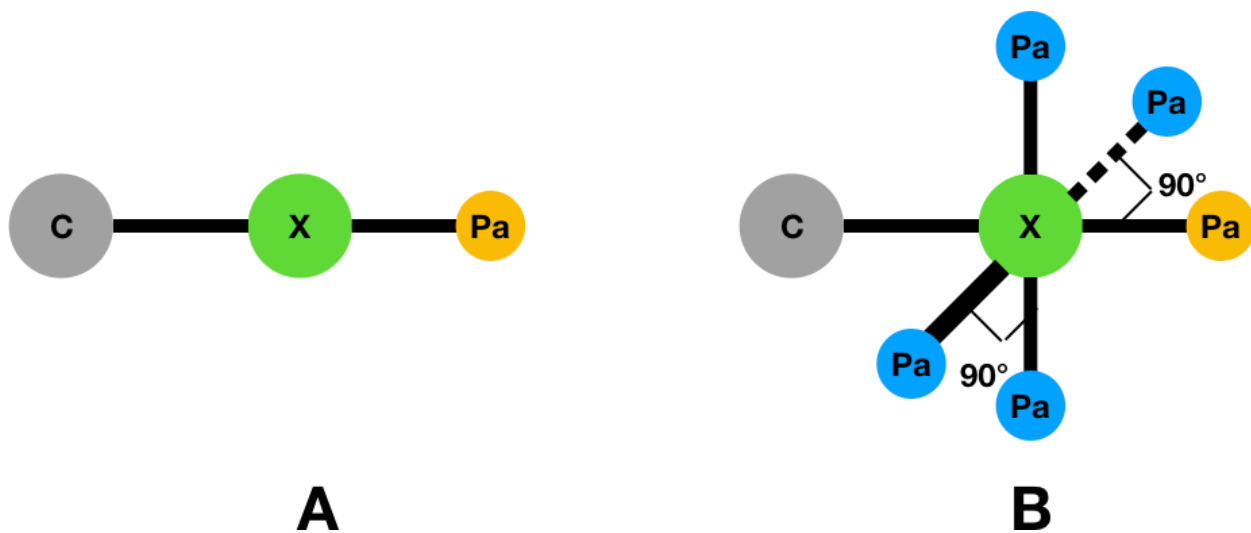
Figure 1. Chlorinated (A÷L) or brominated (M) model systems chosen to study halogen bond with the carbonyl oxygen of the capped alanine through the Pseudo-atom(s) strategy.

MATERIALS AND METHODS

The structures of eleven halogenated molecules complexed by capped alanine (see Figure 1), taken as benchmark in the MM energy minimization, were obtained by quantum mechanical geometry optimizations using the software Gaussian 09.⁴⁶ In particular, for all the complexes DFT (BSSE free) calculations were performed by using the M06-2X⁴⁴ functional, an hybrid meta-GGA functional that is highly recommended to study halogen bonding and other non-covalent interactions;²⁵ the 6-311++G(d,p) basis set has been used for all atoms in the complexes.

The partial atomic charges for the ligands were derived exploiting the RESP procedure, as implemented in AMBER11, on the ESP as obtained by RHF/6-31G(d,p) Gaussian09

1
2
3
4 calculations. To evaluate the goodness of the RESP fitting to the RHF electrostatic potential, the
5
6 relative root mean square error (RRMS) has been monitored. Two different dispositions of
7
8 pseudo-atoms have been then tested (see Figure 2): in the first one ('1Pa' disposition, A), a
9
10 single pseudo-atom (Pa) has been added along the extension of the D–X bond, pointing towards
11
12 the XB acceptor, and several X–Pa distances have been tested as described in the Results section;
13
14 the XB acceptor, and several X–Pa distances have been tested as described in the Results section;
15
16 in the second one ('5Pa' disposition, B), five Pa's were added to the halogenated ligands: one in
17
18 the same position as in '1Pa' and the other four along two directions, orthogonal each other and
19
20 perpendicular to the D–X bond, so that each of these four Pa's is the vertex of a square with the
21
22 chlorine atom in the center. Also in this case several X–Pa distances have been tested. To
23
24 parametrize the halogenated molecules the GAFF^{45,47} force fields have been adopted, while for
25
26 the capped alanine the ff99SBildn force field has been employed.⁴⁸
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51



52 **Figure 2.** Dispositions of pseudo-atoms adopted to describe the anisotropy of electron density
53
54 around the halogen atoms (X, Pa and C stand for halogen, pseudo-atom and carbon,
55
56 respectively). A: '1Pa' disposition where one Pa is used to simulate the charge of the σ -hole. B:
57
58
59
60
61
62
63
64
65

1
2
3
4 ‘5Pa’ disposition, where one Pa (in yellow) is used to simulate the charge of the σ -hole and the
5
6 other four surrounding the halogen are used to enhance the negative belt.
7
8
9

10
11
12 In order to better reproduce the DFT geometrical parameters (interaction distances and angles)
13 of halogen bonding, different values of the van der Waals radius parameter for the chlorine atom
14 have been tested in all the analyzed systems, as described in the Results section, while the well
15 energy depth (the energy minimum for the vdW interaction) has been kept fixed to 0.265 kcal
16 mol⁻¹. Such tests showed that a lowering of the chlorine vdW radius parameter improves the
17 description of the XB geometrical parameters and it is also consistent with the variation of the
18 non-bonded parameters for bromine and iodine passing from the original Amber force field⁴⁰ to
19 the more recent one.
20
21
22
23
24
25
26
27
28
29
30
31

32 The ACPYPE utility has been used to convert the AMBER input files into the Gromacs ones.
33 MM energy minimizations were then performed with Gromacs 5.0.7⁴² package on all complexes
34 either in the native form and after pseudo-atoms insertion for both Pa(s) dispositions. We
35 adopted 10000 steps of steepest descendent minimization, with 10 steps of conjugate gradient
36 minimization for every step of steepest minimization.
37
38
39
40
41
42
43

44 To compare the XB geometrical parameters as obtained from classical energy minimization
45 (em) using the pseudo-atom(s) and DFT approaches, the percentage errors (S_{distance} and S_{angle}), as
46 defined in equations (1) and (2), have been used.
47
48
49
50

$$S_{\text{distance}} = \frac{|d_{em} - d_{DFT}|}{d_{DFT}} \quad (1)$$

$$S_{\text{angle}} = \frac{|ang_{em} - ang_{DFT}|}{ang_{DFT}} \quad (2)$$

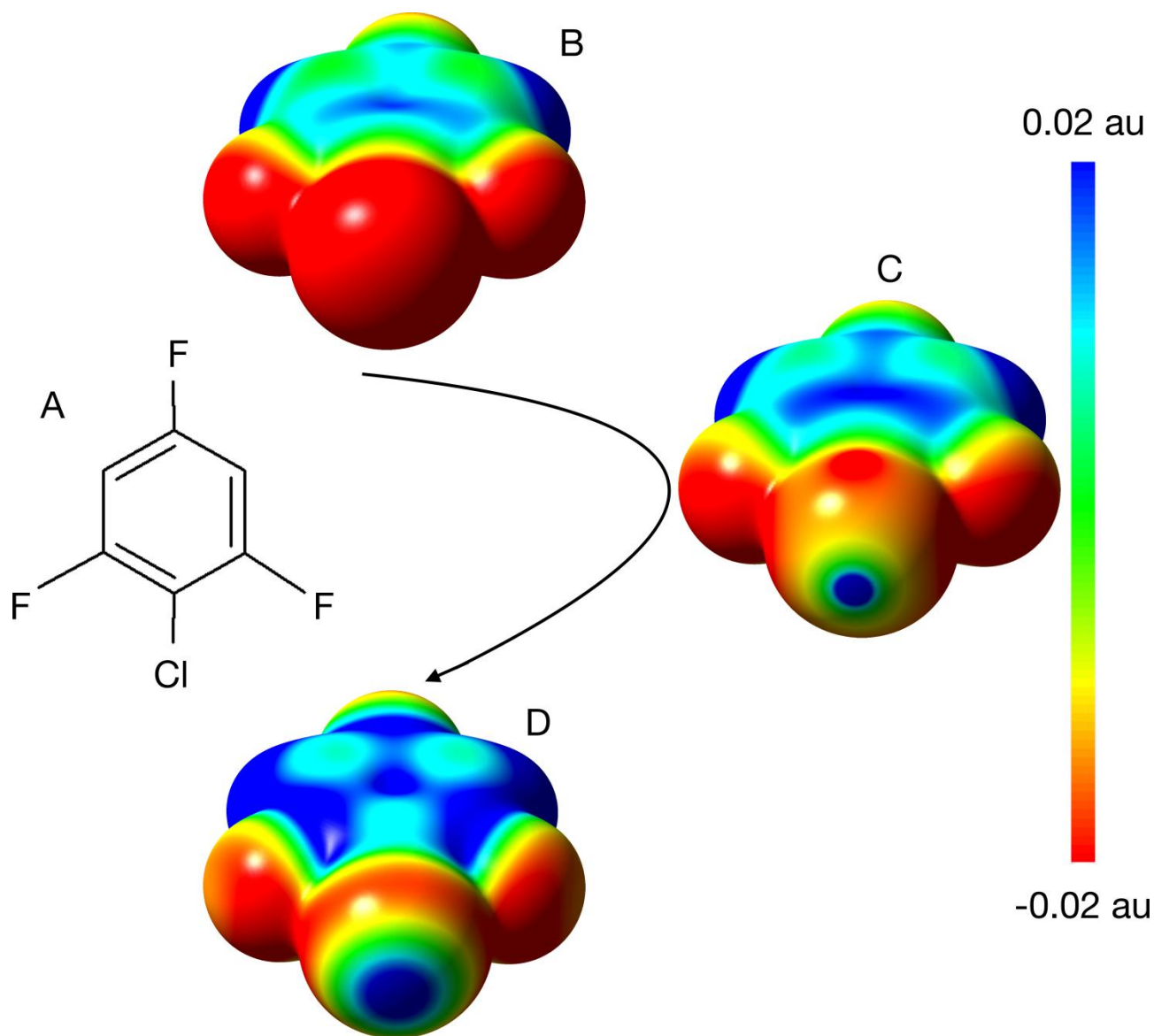
1
2
3
4 where d_{em} and ang_{em} are the interaction distances and angles, respectively, as obtained from the
5
6 classical approach, while d_{DFT} and ang_{DFT} are the corresponding parameters as obtained from
7
8 M06-2X/6-311++G(d,p) geometry optimization.
9

13 14 **RESULTS AND DISCUSSION**

15
16 Halogen bonding established in ten chlorinated model systems with capped alanine (see Figure
17
18 1) has been studied by means of Molecular Mechanics energy minimization with the purpose of
19
20 developing a new strategy to be used in classical force fields methods, owing to the failure of the
21
22 previously reported pseudo-atom methods³⁶⁻³⁸ in describing chlorine-based XB. The new
23
24 approach is based on the use of a set of pseudo-atoms, Pa's, in place of a single one as previously
25
26 proposed.³⁶⁻³⁸ The model systems were properly chosen to simulate halogenated ligands in
27
28 protein-ligand complexes, while the capped alanine has been used because it is a typical XB
29
30 acceptor in protein-ligand systems, as underlined by a PBD survey.¹⁸ As benchmark for testing
31
32 our model systems, the BSSE-free M06-2X/6-311++G(d,p) optimized X...O distances and C-
33
34 X...O angles have been used. Several Pa's models as obtained by varying the number of Pa's and
35
36 their dispositions, their distances from the halogen and the vdW radius of the halogen have been
37
38 tested. The case of no Pa has been as well taken into account.
39
40
41
42
43
44

45
46 The disposition of the Pa's around the halogen is shown in Figure 2. While in the previously
47
48 proposed '1Pa' strategy (Figure 2A) a single Pa was used to reproduce the positive charge of the
49
50 σ -hole, in the present '5Pa' scheme (Figure 2B) four additional Pa's are added to describe the
51
52 negative belt around the halogen atom. As it can be observed from the electrostatic potential
53
54 maps shown in Figure 3, the full set of extra-point charges allows to reproduce the anisotropy of
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 the quantum mechanical ESP by using classical methods, in contrast to the case in which no
5
6 pseudo-atom has been introduced.
7
8
9



51 **Figure 3.** Electrostatic potential mapped onto the electron density surface ($0.002 \text{ electrons au}^{-3}$)
52
53 of 1-chloro-2,4,6-trifluorobenzene (A), calculated at RHF/6-31G(d) level of theory (D) or using
54
55 the RESP procedure without pseudo-atom(s) (B) and after five (C) pseudo-atoms addition at 1.8
56
57 Å from the chlorine atom.
58
59
60
61
62
63
64
65

1
2
3
4
5
6
7 The first important result of the present investigation is that for all the chlorinated systems the
8
9 introduction of one or more Pa's allows to describe the XB interaction through classical energy
10
11 minimization calculations, while the halogen bonding interaction is completely lost (meaningless
12
13 geometrical parameters have been observed) if no extra-point charge is included in the model
14
15 (see Table 2 for results obtained on chlorobenzene using, as an example, 1.748 Å as vdW radius
16
17 for chlorine and 1.8 Å as Cl–Pa(s) distance). However, the inclusion of a single Pa in most cases
18
19 does not allow to correctly reproduce the DFT geometrical parameters for the XB interaction
20
21 (see Figures 4 and S1-S10, and Tables S1 and S2), unlike what previously reported for
22
23 brominated and iodinated systems.³⁸ On the other hand, the '5Pa' strategy brings an important
24
25 improvement in the chlorine XB description, as denoted by a lowering of the percentage
26
27 geometrical errors with respect to DFT results. The fact that the simpler '1Pa' approach is not
28
29 adequate in the case of chlorine could be probably related to the weaker interactions established
30
31 by this atom with respect to those with the heavier halogens, which is a consequence of a less
32
33 pronounced σ -hole. Because of this, the anisotropy of the electrostatic potential around the
34
35 chlorine atom needs to be described in a finer way, which can be realized by introducing more
36
37 than one Pa.
38
39
40
41
42
43
44

45 According to Hobza et al.,³⁶ the optimum distance of the pseudo-atom (explicit sigma hole,
46
47 ESH, in their model) from the halogen should be shorter than the halogen vdW parameter to
48
49 avoid numerical instabilities in simulations. On the other hand, it has been noted that placing the
50
51 Pa too close to the halogen could lead to an overestimation of the extra-point charge. For this
52
53 reason, we decided to test the disposition of the extra-point charges around the halogen atom by
54
55 varying both the X–Pa distance and the chlorine vdW radius, in order to get the better
56
57
58
59
60
61
62
63
64
65

1
2
3
4 compromise for the choice of the X–Pa distance. To this aim, three different X–Pa distances (1.4,
5
6 1.6 and 2.0 Å) have been considered in both strategies (1Pa and 5Pa) for each vdW parameter
7
8 used. The chlorine GAFF vdW parameter has been varied from 1.95 Å, that is the default value,
9
10 to 1.748 and 1.648 Å.
11
12
13

14 As shown in Tables S1, S2 and in Figures S1 through S10, it was not possible to highlight any
15
16 trends in the behavior of the S_{distance} and S_{angle} percentage errors upon variation of the X–Pa
17
18 distances, denoting some insensitiveness of the XB geometry on this parameter in the range here
19
20 considered. On the other hand, concerning the chlorine GAFF VdW parameter, it can be inferred
21
22 that the default vdW radius of chlorine is not suitable for describing this interaction, because,
23
24 almost always, it gives the worst S_{distance} and S_{angle} errors with respect to the others two tested
25
26 values (1.748 and 1.648 Å). This is true for either ‘1Pa’ and ‘5Pa’ strategies but it is more
27
28 evident for the latter one, probably because in the former case the vdW parameter effect is
29
30 “masked” by the greater error introduced by the coarse description of the electrostatic potential
31
32 given by the ‘1Pa’ strategy. The necessity to reduce the vdW radius with respect to the original
33
34 implemented value, in order to correctly describe XB, was already evidenced for bromine and
35
36 iodine, whose vdW radii have been in fact modified from 2.22 to 2.02 Å (Br) and from 2.35 to
37
38 2.15 Å (I) passing from the original⁴⁵ to the improved version of GAFF.⁴⁵ Both of the new
39
40 shorter GAFF vdW radii here tested for chlorine allow to improve the results obtained for all the
41
42 model systems studied, though no systematic differences are observed passing from 1.748 to
43
44 1.648 Å. This conclusion doesn’t claim to be a rigorous re-parametrization of the chlorine GAFF
45
46 vdW radius, but should be view as an indication that this parameter should be revised in future
47
48 GAFF force field versions.
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 It is to be noted that the MM minimization, while providing positive charge, as expected, for
5
6 the ‘ σ -hole’ Pa, resulted in either negative charges for the ‘lateral’ Pa’s in all chlorinated
7
8 derivatives or positive charges for bromobenzene, which was included in our investigation as a
9
10 check of our procedure (see Tables 3 and 4). This observation is in good agreement with the
11
12 results of a charge density investigation by Espinosa et al.,⁵⁰ who explored the differences
13
14 between the chlorine and bromine σ -hole “architectures” based on analysis of the Laplacian of
15
16 electron density, $\nabla^2\rho(r)$, for a series of halogenated molecules. This analysis clearly shows that
17
18 for chlorine it is possible to observe an extended valence shell charge concentration (VSSC)
19
20 region, containing either the charge concentration (CC) and charge depletion (CD) sites. On the
21
22 other side, for bromine a far more reduced VSSC region is observed, containing only the CC
23
24 sites, while the CD sites belong to a large region of positive $\nabla^2\rho(r)$ (which corresponds to a large
25
26 portion of positive electrostatic potential) surrounding the bromine atom. This topology of $\nabla^2\rho(r)$
27
28 is in agreement with the positive values of the four lateral extra-point charges in the case of
29
30 bromobenzene. Moreover, it can explain also the reason why the ‘5Pa’ disposition is necessary
31
32 for correctly describe chlorine halogen bonding: using only one extra-point charge linked to
33
34 chlorine in the position of the σ -hole could induce an underestimation of the electrostatic charge
35
36 related to the σ -hole (because of the extended VSSC region). By introducing the four lateral
37
38 Pa(s), describing the “negative belt” of the halogen’s ESP, the charge of the central Pa fits more
39
40 accurately the real σ -hole charge derived from QM calculations.
41
42
43
44
45
46
47
48
49
50

51
52 The inadequacy of using only one extra-point charge to describe the anisotropy of the
53
54 electrostatic potential around halogen atoms was pointed out also by Hage et al.³⁵ In particular
55
56 their observation on the requirement of the quadrupolar contribution to reproduce
57
58
59
60
61
62
63
64
65

1
2
3
4 thermodynamic properties of halogen bonding (in particular for lighter halogens) by using hybrid
5
6 (PC/MTP) approaches would help corroborate the ‘5Pa’ model here proposed.
7
8
9

10
11
12
13
14 **Table 2.** Halogen bonding interaction distances and angles for the chlorobenzene-capped alanine
15
16 dimer, computed with molecular mechanics using 0, 1 and 5 pseudo-atoms and with M06-2X/6-
17
18 311++G(d,p) geometry optimization.^a
19
20
21

Chlorobenzene	Distance/Å	Angle/°
No Pseudo-atom	Lost	Lost
1 Pseudo-atom	3.02	169.97
5 Pseudo-atoms	3.24	161.14
M06-2X/6-311++G(d,p)	3.17	160.11

22
23
24
25
26
27
28
29
30
31
32
33 ^aMM calculations performed using 1.748 Å as VdW parameter for the chlorine atom and 1.8 Å
34
35 as Pa(s) distances from chlorine.
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

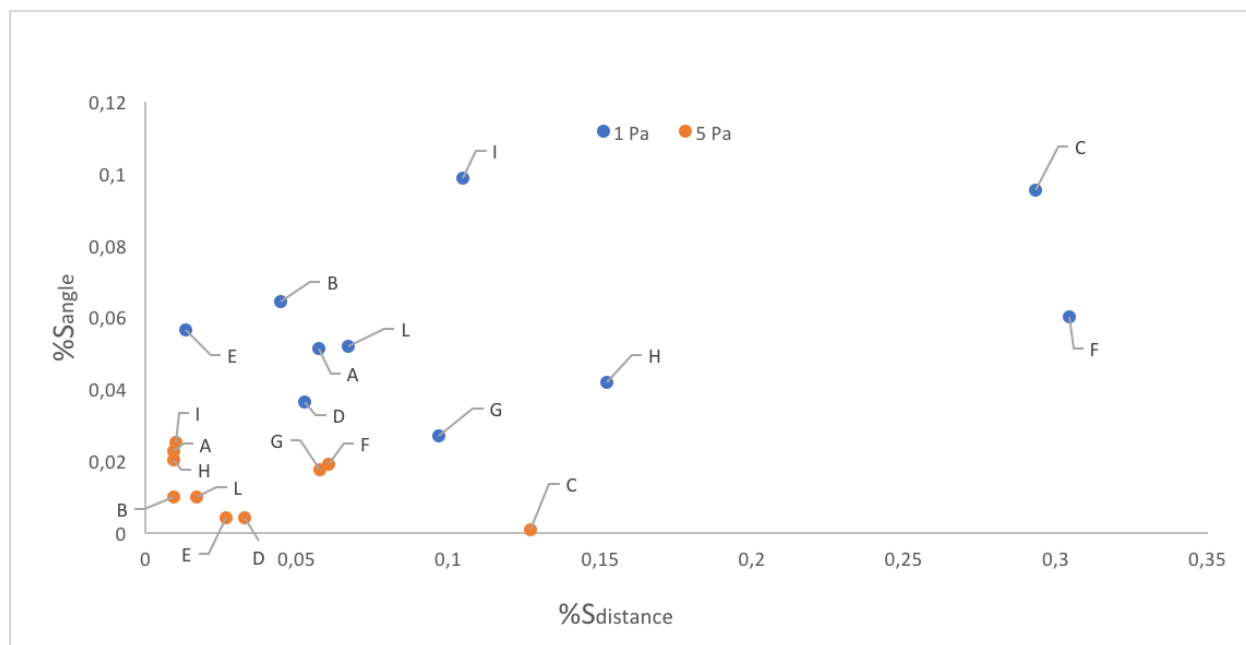


Figure 4. Percentage errors on halogen bonding distances (x-axis) and angles (y-axis), for the ten chlorinated systems studied in this work. ‘5Pa’ strategy in red, ‘1Pa’ strategy in blue (calculations performed using 1.748 Å as vdW parameter for the chlorine atom and 2.0 Å as Pa(s) distances from the central halogen).

Table 3. Selected RESP derived atomic point charges (in amu) in chlorobenzene: Carbon linked to chlorine, chlorine, pseudo-atom representing the σ -hole and the four lateral pseudo-atoms. 1.8 Å as Pa(s) distances from chlorine have been used.

	C	Cl	‘ σ -hole’ Pa	Pa	Pa	Pa	Pa
No Pa	-0.01	-0.13	-	-	-	-	-
1 Pa	0.02	-0.15	0.01	-	-	-	-
5 Pa	0.30	-0.23	0.04	-0.01	-0.01	-0.01	-0.01

Table 4. Selected RESP derived atomic point charges (in amu) in bromobenzene: Carbon linked to bromine, bromine, pseudo-atom representing the σ -hole and the four lateral pseudo-atoms. 1.8 Å as Pa(s) distances from chlorine have been used.

	C	Br	' σ -hole' Pa	Pa	Pa	Pa	Pa
No Pa	-0.12	-0.09	-	-	-	-	-
1 Pa	0.20	-0.30	0.06	-	-	-	-
5 Pa	0.15	-0.32	0.03	0.02	0.02	0.02	0.02

For all studied systems, a decreasing of the RRMS error related to the RESP procedure has been observed passing from no to one and then to five Pa's. As pointed out in the previous pseudo-atoms studies for halogen bond,³⁶⁻³⁸ the introduction of an extra point charge linked to the halogen atom leads to an improvement of the description of the electrostatic potential around the halogens. The introduction of more than one Pa reduces only slightly the RRMS error, much less than what obtained passing from zero to one Pa. However, it is well known that small changes in the molecular electrostatic potential may involve even large variations in the chemistry of non-covalent interactions. Accordingly, the great improvement in the geometrical parameters of halogen bonded complexes observed upon insertion of more than one Pa derives from small modifications of the ESP, though they result in only limited improvement in the values the RRMS associated to the RESP procedure.

CONCLUSIONS

A series of chlorinated model systems have been studied by molecular mechanics energy minimization exploiting a set of extra-point charges to simulate the σ -hole and the negative belt

1
2
3
4 characterizing the electron density distribution around the chlorine atom. Differently from the
5
6 previously reported pseudo-atom ('1Pa') approach,³⁶⁻³⁸ the '5Pa' strategy here proposed is
7
8 demonstrated to well reproduce quantum mechanical geometrical parameters for halogen
9
10 bonding interaction of the model systems with the carbonyl oxygen of the capped alanine. The
11
12 model systems have been chosen to either reproduce real systems that are likely to be
13
14 encountered in PDB and test the strategy to a rather extended range of XB strengths.
15
16
17

18
19 We found out that the MM description of chlorine XB has been greatly improved by the
20
21 introduction of more than one Pa's in all the chlorinated model systems here investigated. This
22
23 improvement is evident from the lowering of the percentage errors on the interaction distances
24
25 and angles, as referred to benchmark M06-2X/6-311++G(d,p) geometry optimizations, when
26
27 going from the '1Pa' to the '5Pa' methodology. The same improvement was not observed in the
28
29 case of bromobenzene, suggesting the correctness of the '1Pa' strategy in the description of
30
31 halogen bonding when bromine atoms are involved.
32
33
34

35
36 Moreover, it has been shown (through testing of different vdW radius values) that a re-
37
38 parametrization of the chlorine GAFF van der Waals radius should be taken into account to
39
40 obtain a proper halogen bonding description.
41
42

43 Further development of this methodology will comprehend either testing our Pa's strategy on
44
45 other model compounds (including different halogen bond acceptor) and performing molecular
46
47 dynamic simulations on condensed phase systems (like complexes between proteins and
48
49 chlorinated ligands).
50
51
52

53
54
55 ASSOCIATED CONTENT
56

57
58
59 **Supporting Information.**
60
61
62
63
64
65

1
2
3
4 Geometrical parameters (Tables S1÷S3) and two-dimensional plots of the S_{angle} and S_{distance}
5 errors for all the studied XB systems (Figures S1÷S10).
6
7

8 9 **Author Contributions**

10 The manuscript was written through contributions of all authors. All authors have given approval
11 to the final version of the manuscript.
12
13
14

15 16 17 **ACKNOWLEDGEMENTS**

18 The Fondazione Banca del Monte di Lombardia is fully acknowledged for financial support.
19
20

21 22 **ABBREVIATIONS**

23 QM/MM, quantum mechanics/molecular mechanics; XB, halogen bonding; X, halogen; D, donor
24 group; A, acceptor group; ESP, electrostatic potential; HB, hydrogen bonding; PDB, protein data
25 bank; DFT, density functional theory; RESP, restrained electrostatic potential); Pa, pseudo-atom;
26 Pa's, pseudo-atoms; MM, molecular mechanics; GAFF, general AMBER force field; BSSE,
27 basis set superposition error; GGA, general gradient approximation; RRMS, root mean square
28 error; VdW, Van der Waals; VSCC, valence shell charge concentration; CC, charge
29 concentration; CD, charge depletion; QM, quantum mechanics; PC/MTP, point
30 charge/multipole.
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

48 49 **REFERENCES**

- 50
51 1. Clark, T; Hennemann, M; Murray, J. S.; Politzer, P. *J. Mol. Mod.*, **2007**, 13, 291-296.
52
53
54 2. Politzer, P.; Murray, J. S. *Crystals* **2017**, 7, 212.
55
56
57 3. Riley, K. E.; Murray, J. S.; Politzer, P; Concha, M. C.; Hobza, P. *J. Chem. Theory*
58 *Comput.* **2009**, 5, 155.
59
60
61
62
63
64
65

- 1
2
3
4 4. Riley, K. E.; Murray, J. S.; Fanfrlik, J.; Řezáč, J.; Solà, R. J.; Concha, M. C.; Ramos, F.
5
6 M.; Politzer, P. *J. Mol. Model.* **2011**, 17, 3309.
7
8
9
10 5. Lu, Y.; Zou, J.; Wang, Y.; Jiang, Y.; Yu, Q. *J. Phys. Chem. A*, **2007**, 111, 42, 10781-
11
12 10788.
13
14
15 6. Forni, A.; Lucenti, E.; Botta, C.; Cariati, E. *J. Mater. Chem. C* **2018**, 6, 4603-4626.
16
17
18
19 7. Cariati, E.; Cavallo, G.; Forni, A.; Leem, G.; Metrangolo, P.; Meyer, F.; Pilati, T.;
20
21 Resnati, G.; Righetto, S.; Terraneo, G.; Tordin, E. *Cryst. Growth Des.* **2011**, 11, 12,
22
23 5642–5648.
24
25
26
27 8. Thallapally, P. K. ; Desiraju, G. R. ; Bagieu-Beucher, M. ; Masse, R. ; Bourgogne, C. ;
28
29 Nicoud, J. *Chem. Commun.* **2002**, 1052–1053.
30
31
32
33 9. Christodoulou, M. S.; Zunino, F.; Zuco, V.; Borrelli, S.; Comi, D.; Fontana, G.;
34
35 Martinelli, M.; Lorens, J. B.; Evensen, L.; Sironi, M.; Pieraccini, S.; Via, L. D.; Gia, O.
36
37 M.; Passarella, D. *Chem. Med. Chem.*, **2012**, 7, 2134-2143.
38
39
40
41 10. Voth, R.; Hays, F. A.; Ho, P. S. *Proc. Natl. Acad. Sci., U. S. A.* **2007**, 104, 15,
42
43 6188–6193.
44
45
46
47 11. Hardegger, L. A.; Kuhn, B.; Spinnler, B.; Anselm, L.; Ecabert, R.; Stihle, M.; Gsell, B.;
48
49 Thoma, R.; Diez, J.; Benz, J.; Plancher, J.; Hartmann, G.; Banner, D. W.; Haap, W.;
50
51 Diederich, F.; *Angewandte Chemie*, **2010**, 50, 1, 314-318.
52
53
54
55 12. Christodoulou, M. S.; Calogero, F.; Baumann, M.; García-Argáez, A. N.; Pieraccini, S.;
56
57 Sironi, M.; Dapiaggi, F.; Bucci, R.; Broggin, G.; Gazzola, S.; Liekens, S.; Silvani, A.;
58
59
60
61
62
63
64
65

- 1
2
3
4 Lahtela-Kakkonen, M.; Martinet, N.; Nonell-Canals, A.; Santamaría-Navarro, E.;
5
6 Baxendale, I. R.; Dalla Via, L.; Passarella, D. *Eur. J. Med. Chem.* **2015**, 92, 766-775.
7
8
9
10 13. Ho, P. S. *Future Med. Chem.* **2017**, 9, 637-640.
11
12
13 14. Christodoulou, M. S.; Sacchetti, A.; Ronchetti V.; Caufin, S. A.; Silvani, A.; Lesma G.;
14
15 Fontana, G.; Minicone, F.; Riva, B.; Ventura, M.; Lahtela-Kakkonen, M; Jarho, E.; Zuco,
16
17 V.; Zunino, F.; Martinet, N.; Dapiaggi, F.; Pieraccini, S.; Sironi, M.; Dalla Via, L.; Gia,
18
19 O. M.; Passarella, D. *Bioorganic Med. Chem.*, **2013**, 21, 6920-6928.
20
21
22
23 15. Rowe, R. K.; Ho, P. S. *Acta Cryst. B*, **2017**, B73, 255-264.
24
25
26 16. Carlsson, A. C.; Scholfield, M. R.; Rowe, R. K.; Ford, M. C.; Alexander, A. T.; Mehl, R.
27
28 A.; Ho, P. S. *Biochemistry*, **2018**, 57, 4135-4147.
29
30
31
32 17. Hardegger, L. A.; Kuhn B.; Spinnler, B.; Anselm, L.; Ecabert, R.; Stihle, M.; Gsell, B.;
33
34 Thoma, R.; Diez, J.; Benz, J.; Plancher, J. M.; Hartmann, G.; Isshiki, Y.; Morikami, K.;
35
36 Shimma, N.; Haap, W.; Banner, D. W.; Diederich, F. *Chem. Med. Chem.*, **2016**, 6, 2048-
37
38 2054.
39
40
41
42 18. Lu, Y.; Shi T.; Wang Y.; Yang, H.; Yan, X.; Luo, X.; Jiang, H.; Zhu, W. *J. Med. Chem.*
43
44 **2009**, 52, 2854-2862.
45
46
47
48 19. Forni, A.; Pieraccini, S.; Rendine, S.; Sironi, M. *J. Comp. Chem.* **2014**, 35, 386.
49
50
51
52 20. Forni, A.; Pieraccini, S.; Rendine, S.; Gabas, F.; Sironi, M. *ChemPhysChem* **2012**, 13,
53
54 4224.
55
56
57 21. Forni, A.; Pieraccini, S.; Franchini, D.; Sironi, M. *J. Phys. Chem. A* **2016**, 120, 9071.
58
59
60
61
62
63
64
65

- 1
2
3
4 22. Kozuch, S.; Martin, J. M. L. *J. Chem. Theory Comput.*, **2013**, 9, 1918.
5
6
7
8 23. Lu, Y.; Li, H.; Zhu, X.; Zhu, W.; Liu, H. *J. Phys. Chem. A* **2011**, 115, 4467.
9
10
11 24. Lu, Y.; Li, H.; Zhu, X.; Liu, H.; Zhu, W. *Int. J. Quantum Chem.* **2012**, 112, 1421.
12
13
14 25. Forni, A.; Rendine, S.; Pieraccini, S.; Sironi, M. *J. Mol. Graph. Model.* **2012**, 38, 31.
15
16
17 26. Cornell, W. D.; Cieplak, P.; Bayly C. I.; Kolman, P. A. *J. Am. Chem. Soc.* **1993**, 115,
18
19 9620–9631.
20
21
22 27. Ford, M. C.; Ho, P. S. *J. Med. Chem.*, **2017**, 60, 8681-8690.
23
24
25
26 28. Hage, K. E.; Bereau, T.; Jakobsen, S.; Meuwly, M. *J. Chem. Theory Comput.* **2016**, 12, 7,
27
28 3008-3019.
29
30
31
32
33 29. Jorgensen, W. L.; Schyman, P. *J. Chem. Theory Comput.* **2012**, 8, 10, 3895-3901.
34
35
36
37
38 30. Harder, E.; Damm, W.; Maple, J.; Wu, C.; Reboul, M.; Xiang, J. Y.; Wang, L.; Lupyan,
39
40 D.; Dahlgren, M. K.; Knight, J. L.; Kaus, J. W.; Cerutti, D. S.; Krilov, G.; Jorgensen, W.
41
42 L.; Abel, R.; Friesner, R. A. *J. Chem. Theory Comput* **2016** 12, 1, 281-296.
43
44
45
46
47 31. Mu, X.; Wang, Q.; Wang, L. P.; Fried, S. D.; Piquemal, J. P.; Dalby, K. N.; Ren, P. *J.*
48
49 *Phys. Chem. B* **2014** 118, 24, 6456-6465.
50
51
52
53
54 32. Gutiérrez, I. S.; Lin, F. Y.; Vanommeslaeghe, K.; Lemkul, J. A.; Armacost, K. A.;
55
56 Brooks, C. L.; MacKerell Jr, A. D. *Bioorg. Med. Chem.* **2016**, 24, 4812-4825.
57
58
59
60
61
62
63
64
65

- 1
2
3
4 33. Ho, P. S. In: *Metrangolo P.; Resnati G., Halogen Bonding I. Topics in Current*
5
6
7 *Chemistry, vol. 358., 2015*, Springer, Cham. (Pages 241-276).
8
9
10 34. Scholfield, M. R.; Ford, M. C.; Vander Zanden, C. M.; Billman, M. M.; Ho, P. S.; Rappé,
11
12 A. K. *J. Phys. Chem. B* **2015**, 119, 29, 9140-9149.
13
14
15 35. Santos, L. A.; da Cunha, E. F. F.; Ramalho, T. C. *J. Phys. Chem. A* **2017**, 121, 2442
16
17
18
19 36. Kolár, M.; Hobza, P. *J. Chem. Theory Comput.* **2012**, 8, 4, 1325-1333.
20
21
22 37. Ibrahim, M. A. A. *J. Chem. Theory Comput.* **2011**, 32, 12, 2564-2574.
23
24
25
26 38. Rendine, S.; Pieraccini, S.; Forni, A.; Sironi, M. *Phys. Chem. Chem. Phys.* **2011**, 13, 43,
27
28 19508-19516.
29
30
31 39. Auffinger, P.; Hays, F. A.; Westhof, E.; Ho, P. S. *PNAS* **2004**, 101, 48, 16789-16794.
32
33
34
35 40. Eskandari, K.; Lesani, M. *Chem. Eur. J.* **2015**, 21, 1-9.
36
37
38 41. Lin, F. Y.; MacKerrell, A. D.; *J. Chem. Theory Comput.* **2018**, 14, 2, 1083-1098.
39
40
41 42. Lin, F. Y.; MacKerrell, A. D. *J Phys. Chem. B* 2017, 121, 28, 6813-6821.
42
43
44 43. Wilcken, R.; Zimmermann, M. O.; Lange, A.; Joerger, A. C.; Boeckler, F. M. **2013**, *J.*
45
46 *Med. Chem.* **2013**, 56, 4, 1363-1388.
47
48
49
50 44. Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, 120, 1-3, 215-241.
51
52
53
54 45. Özpınar, G. A.; Peukert, W.; Clark T. *J Mol Model* **2010**, 16, 1427–1440.
55
56
57 46. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman,
58
59 J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato,
60
61
62
63
64
65

1
2
3
4 M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.;
5
6 Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.;
7
8
9 Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.;
10
11 Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.;
12
13 Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.;
14
15 Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.;
16
17 Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.;
18
19 Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.;
20
21 Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels,
22
23 A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian 09,
24
25 Revision D.01 Gaussian, Inc., Wallingford CT, 2013.
26
27
28
29
30
31

32 47. Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman P. A.; Case, D. A. *J. Comput. Chem.*

33
34 **2004**, 25, 1157–1174.
35

36 48. Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror R.O.; Shaw,

37
38 D. E. *Proteins* **2010**, 78, 1950.
39

40 49. Abraham, M. J.; Murtola, T.; Schulz, R.; Pall, S.; Smith, J. C.; Hess, B.; Lindahl, E.

41
42
43 *SoftwareX*, **2015**, 1-2 19-25.
44

45 50. Brezgunova, M. E.; Aubert, E.; Dahaoui, S.; Fertey, P.; Lebègue, S.; Jelsch, C.; Ángyán,

46
47
48 G. J.; Espinosa, E. *Cryst. Growth Des.* **2012**, 12, 11, 5373-5386.
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65