Halogen bonding at the wet interface of an amyloid peptide structure

Andrea Pizzi,[a] *Nicola Demitri,*[b] *Giancarlo Terraneo*,*[a] *and Pierangelo Metrangolo*[a]

[a] Laboratory of Supramolecular and Bio-Nanomaterials (SBNLab), Department of Chemistry, Materials, and Chemical Engineering "Giulio Natta", Politecnico di Milano, Via Luigi Mancinelli 7, Milano I-20131, Italy.

[b] Elettra – Sincrotrone Trieste, S.S. 14 Km 163.5 in Area Science Park, 34149 Basovizza – Trieste, Italy.

ELECTRONIC SUPPLEMENTARY INFORMATION

Supplementary material

CCDC 1574297 and 1574298 contain the supplementary crystallographic data for compounds **DF(Cl)NKF(Cl)** and **DF(I)NKF(I)**. These data can be obtained free of charge from The Cambridge Crystallographic Data Centre via https://www.ccdc.cam.ac.uk/structures.

X-ray diffraction analysis - Structural characterization of DF(Cl)NKF(Cl) and DF(I)NKF(I) peptides

Data collections were performed at the X-ray diffraction beamline (XRD1) of the Elettra Synchrotron, Trieste (Italy)[1S]. The crystals were dipped in perfluoropolyether Fomblin oil (Sigma Aldrich, Saint Louis, USA) and mounted on the goniometer head with kapton loops (MiTeGen, Ithaca, USA). Complete datasets were collected at 100 K (nitrogen stream supplied through an Oxford Cryostream 700 - Oxford Cryosystems Ltd., Oxford, United Kingdom) through the rotating crystal method. Data were acquired using a monochromatic wavelength of 0.800 Å for **DF(Cl)NKF(Cl)** and 0.700 Å for **DF(I)NKF(I)**, on a Pilatus 2M hybrid-pixel area detector (DECTRIS Ltd., Baden-Daettwil, Switzerland). The diffraction data were indexed and integrated using XDS.[2S] For both molecules two different datasets, collected from different crystals randomly oriented, have been merged. Semi-empirical absorption corrections and scaling were performed on datasets, exploiting multiple measures of symmetry-related reflections, using SADABS program.[3S] Crystals appear as very thin colorless needles or plates prone to radiation damage, as previously reported for other halogenated molecules.[4S, 5S] The structures were solved by the dual space algorithm implemented in the SHELXT code.[6S] Fourier analysis and refinement were performed by the full-matrix least-squares methods based on F^2 implemented in SHELXL (Version 2017/1)[7S]. The Coot program was used for modeling.[8S] Anisotropic thermal motion refinement have been used for all atoms with occupancy greater than 50% in **DF(I)NKF(I)**. None of the **DF(Cl)NKF(Cl)** crystals tested diffracted better than ~1.1 Å, and considering radiation damage, the overall dataset resolution is not better than ~1.2 Å. The number of data for the chlorinated model fitting was therefore limited and, to avoid over-refinement, anisotropic thermal motion modeling has been applied only to halogen and backbone oxygen atoms of the peptide (the heaviest atoms in the molecules considered). Geometric and thermal motion parameters restrains (DFIX, DANG, SIMU or ISOR) have been applied on poorly defined fragments. Hydrogen atoms were included at calculated positions with isotropic $U_{factors} = 1.2\ U_{eq}$ or $U_{factors} = 1.5\ U_{eq}$ for protonated amino groups (U_{eq} being the equivalent isotropic thermal factor of the bonded non hydrogen atom). Hydrogen atoms for water molecules have not been included in the refined models since it was not possible to

locate them unambiguously in electron-density peaks of Fourier difference maps (contributions of these missing H atoms are still included in the properties reported in Table 1S). Refined Flack parameters[9S] confirm the expected amino acids Cα configurations. Pictures were prepared using Ortep3[10S], CCDC Mercury[11S] and Pymol[12S] software. Essential crystal and refinement data are reported below (Table 1S).

The two peptides crystallize in the chiral monoclinic *C* 2 space group with one molecule in the asymmetric unit (ASU; Figure S1), with similar lattice parameters and equivalent packing (Figure S2). Cell volume is slightly bigger for the iodinated peptide, as expected from comparison of halogens atomic radius. Superimposition of the two peptides conformations shown an almost perfect match (r.m.s.d. 0.744 Å, Figure S3). Salt bridges links C-terminal carboxylate with primary amine of lysine and protonated N-terminal with aspartate side chain of flanked molecules (shorter $O^{-}\bullet\bullet\bullet N^{+}$ distance is 2.76(3) Å in chlorinated, 2.762(9) Å in iodinated peptide). Peptides adopt an elongated conformation with head-to-head and tail-to-tail non covalent connections through the two kind of ionic interactions identified. Furthermore strong hydrogen bonds give rise to extended parallel β sheets, forming layers of molecules parallel to *bc* plane (shown with alternate carbon colors in figure 2S). Hydrogen bonds involve amidic groups of neighbor peptidic chains. Polar groups on sidechains (N, K and D) and N,C-terminals also cooperate to the stabilization of β sheets (H-bonds details in Table 2S). Both the structures show how the halogens sit packed around atoms of surrounding peptides. The iodine atoms interact with water molecules via halogen bonds while chlorine atoms do not shown any specific interactions with other residues, only very weak hydrogen bonds are detected at sum of vdW radii + 0.20 Å.

Water molecules fill channels parallel to crystallographic *b* axis and are tightly bounded to peptide heteroatoms through hydrogen bonds. Solvent voids represent 25.1% of the cell volume for **DF(Cl)NKF(Cl)** (1112 $Å^3$) and 27.0% of the cell volume for **DF(I)NKF(I)** (1304 $Å^3$)[13S]. Stacked layer of peptides are glued together by bridging water molecules and minor hydrophobic interactions between neighbor halogenated phenyl rings.

The water molecules at the wet interface are held in place by an extended network of hydrogen and halogen bonds. The thermal factors (ADPs) of the water molecules directly in contact with the iodinated residue are less pronounced compared to those of the outer shell. This evidence support the presence of good halogen bonding interaction between iodine atoms and the oxygen atoms of some water molecules. The U(equiv) value, defined as U(equiv) = (1/3) sum~i~[sum~j~(U^ij^

a*~i~ a*~j~ a~i~ a~j~)] where a = the real-space cell lengths and a* = the reciprocal-space cell lengths [14S], for the two oxygen atom involving in the XB are 0.059 Å^2 for O21 and 0.077 Å^2 for O26 where for the other water molecules the average U(equiv) is close to 0.097 Å^2. These values highlight the different "dynamic behavior" of the water molecules in the inner or outer shell. This analysis can be qualitatively done between in the two peptides, the average U(equiv) value for water molecules in the AU for the two structures are 0.083 Å^2 for the iodinated sequence and 0.221 Å^2 for the chlorinated sequence, respectively. This suggests a higher mobility of the water molecules in the **DF(Cl)NKF(Cl)** than in **DF(I)NKF(I)**.

Figure S1. Ellipsoids representation of ASU contents (50% probability) for: A) **DF(Cl)NKF(Cl)** and B) **DF(I)NKF(I)**.

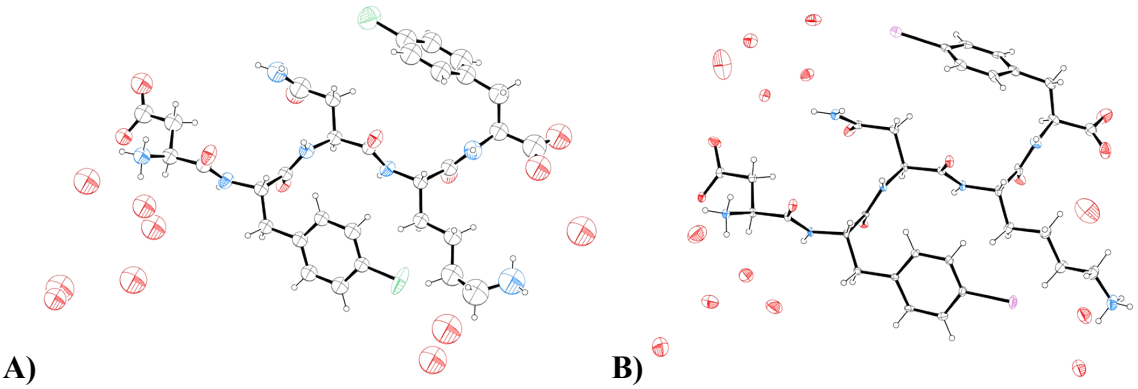


Figure S2. Crystal packing views of **DF(Cl)NKF(Cl)** and **DF(I)NKF(I)** (from top to bottom) along crystallographic *a*, *b* and *c* directions (hydrogens omitted for clarity).

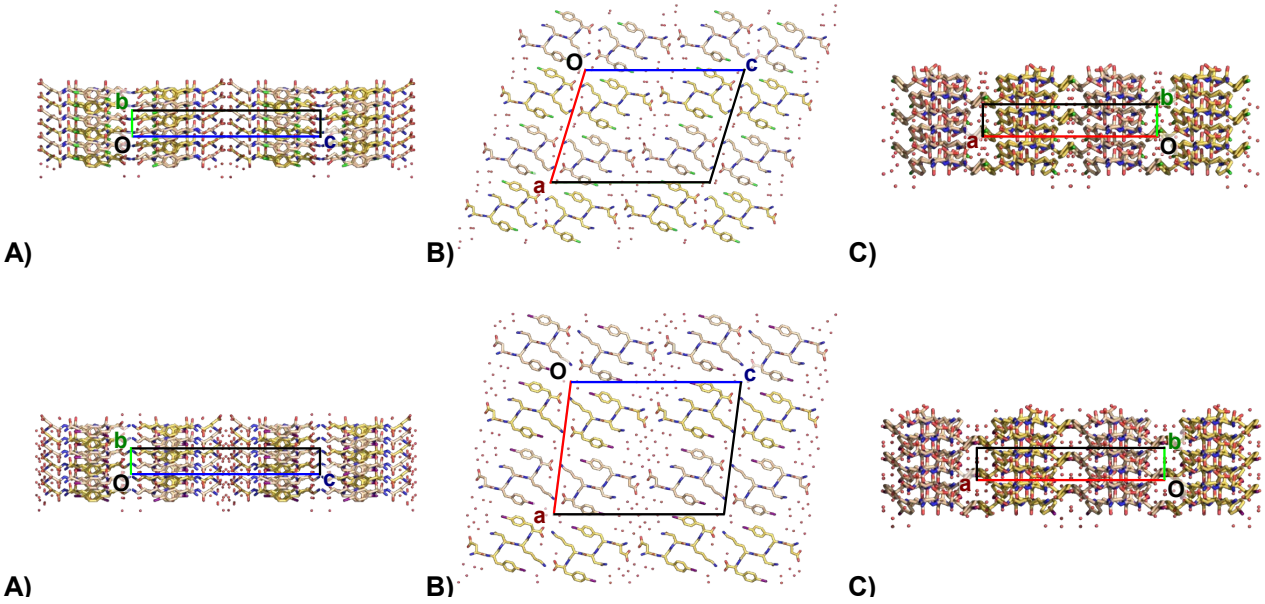


Figure S3. Superimposition of **DF(Cl)NKF(Cl)** and **DF(I)NKF(I)** conformations (light grey and dark grey sticks respectively). Root mean square deviation between overlapped atoms is 0.74 Å.

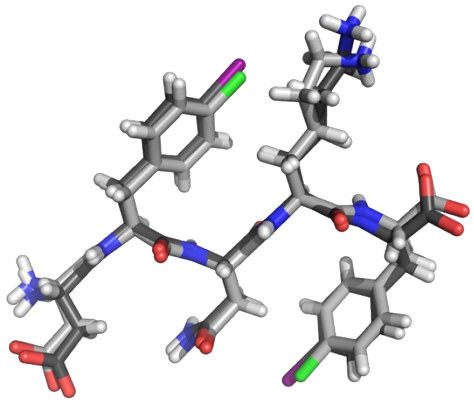
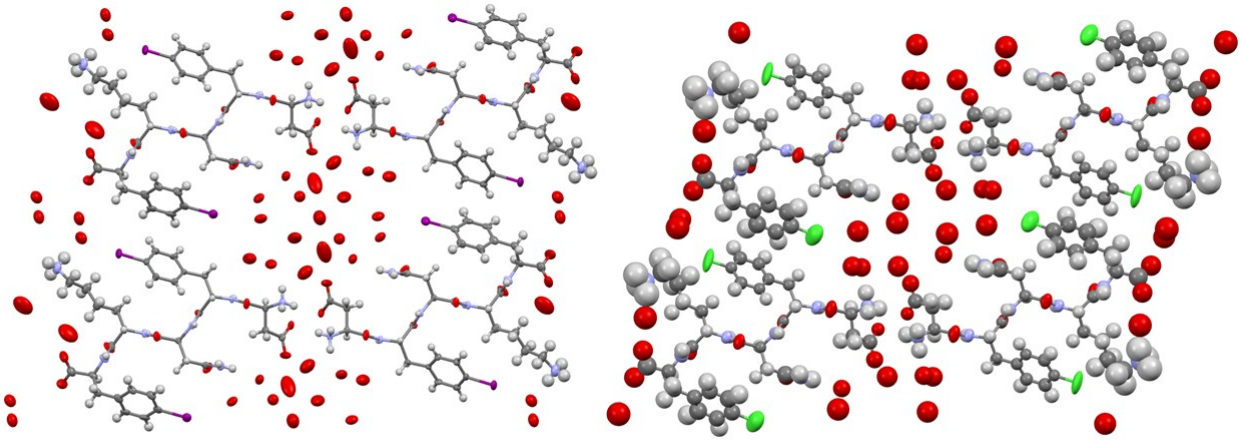


Figure S4. Water content comparison between **DF(I)NKF(I)** and **DF(Cl)NKF(Cl)**. The structures are shown as ellipsoid (50 % probability level) for a direct comparison between the ADPs the water molecules interacting the halogenated residues and those which are not in direct contact.



DF(I)NKF(I)

DF(Cl)NKF(Cl)

Table 1S. Crystallographic data and refinement details for **DF(Cl)NKF(Cl)** and **DF(I)NKF(I)**.

	DF(Cl)NKF(Cl) $\cdot 7H_2O$ $[C_{32}H_{41}Cl_2N_7O_9 \cdot 7H_2O]$	**DF(I)NKF(I)** $\cdot 11.5H_2O$ $[C_{32}H_{41}I_2N_7O_9 \cdot 11.5H_2O]$
CCDC Number	1574297	1574298
Chemical Formula	$C_{32}H_{55}Cl_2N_7O_{16}$	$C_{32}H_{64}I_2N_7O_{20.5}$
Formula weight (g/mol)	864.73	1128.70
Temperature (K)	100(2)	100(2)
Wavelength (Å)	0.800	0.700
Crystal system	Monoclinic	Monoclinic
Space Group	*C* 2	*C* 2
Unit cell dimensions	*a* = 26.685(5) Å	*a* = 26.206(6) Å
	b = 4.809(1) Å	*b* = 4.803(1) Å
	c = 36.125(7) Å	*c* = 35.888(7) Å
	$\alpha = 90°$	$\alpha = 90°$
	$\beta = 107.56(3)°$	$\beta = 97.44(3)°$
	$\gamma = 90°$	$\gamma = 90°$
Volume (Å^3)	4419.8(17)	4820.9(17)
Z	4	4
Density (calculated) (g·cm^{-3})	1.300	1.555
Absorption coefficient (mm^{-1})	0.296	1.311
F(000)	1832	2300
Crystal size (mm^3)	0.09 x 0.01 x 0.01	0.20 x 0.03 x 0.03
Crystal habit	Colorless thin needles	Colorless thin plates
Theta range for data collection	1.72° to 19.30°	1.13° to 29.99°
Resolution (Å)	1.21	0.70
Index ranges	$-20 \le h \le 22$	$-39 \le h \le 37$
	$-3 \le k \le 3$	$-6 \le k \le 6$
	$-29 \le l \le 29$	$-51 \le l \le 51$
Reflections collected	6299	44557
Independent reflections (data with I>2σ(I))	2411 (1373)	13701 (10227)
Data multiplicity (max resltn)	4.12 (2.07)	5.48 (4.45)
I/σ(I) (max resltn)	7.19 (3.17)	10.93 (3.56)
R_{merge} (max resltn)	0.1111 (0.2604)	0.0946 (0.4029)
Data completeness (max resltn)	96.5% (84.6%)	98.8% (96.6%)
Refinement method	Full-matrix least-squares on F^2	Full-matrix least-squares on F^2
Data / restraints / parameters	2411 / 195 / 263	13701 / 7 / 561
Goodness-of-fit on F^2	1.136	1.026
Δ/σ_{max}	0.005	0.001
Final R indices [I>2σ(I)][a]	$R_1 = 0.1420$, $wR_2 = 0.2954$	$R_1 = 0.0714$, $wR_2 = 0.1751$
R indices (all data)[a]	$R_1 = 0.2381$, $wR_2 = 0.3499$	$R_1 = 0.0960$, $wR_2 = 0.1904$
Flack x parameter	-0.27(17)	0.009(13)
Largest diff. peak and hole (e·Å^{-3})	0.519 and -0.407	2.306 and -1.903
R.M.S. deviation from mean (e·Å^{-3})	0.099	0.158

[a] $R_1 = \Sigma \, ||Fo|-|Fc|| \, / \, \Sigma \, |Fo|$, $wR_2 = \{\Sigma \, [w(Fo^2 - Fc^2)^2] \, / \, \Sigma \, [w(Fo^2)^2]\}^{½}$

Table 2S. Geometrical parameters of hydrogen bonds found between peptides in **DF(Cl)NKF(Cl)** and **DF(I)NKF(I)** crystal packing.

DF(Cl)NKF(Cl)				
D-H···A	**d(D-H) (Å)**	**d(H···A) (Å)**	**d(D···A) (Å)**	**<(DHA) (°)**
N-H0B_11•••OD1_11#1	0.91	1.99	2.85(4)	157.2
N-H0C_11•••OD1_11	0.91	2.55	3.10(3)	119.1
N-H0C_11•••OD1_11#2	0.91	1.98	2.76(3)	142
CA-HA_11•••O_11#3	1	2.4	3.20(2)	136.3
N-H0_12•••O_11#3	0.88	2.08	2.94(3)	167.6
CA-HA_12•••O_12#1	1	2.36	3.28(2)	151.6
N-H0_13•••O_12#1	0.88	2.08	2.92(3)	159
CA-HA_13•••O_13#3	1	2.41	3.31(2)	149.8
CB-HB1_13•••OD1_13#1	0.99	2.38	3.29(3)	151.7
ND2-HD2B_13•••OD1_13#1	0.88	2.14	2.97(2)	156.1
N-H0_14•••O_13#3	0.88	1.99	2.86(3)	168.3
CD-HD1_14•••ClZ_12	0.99	2.84	3.43(5)	119.1
CE-HE1_14•••ClZ_12	0.99	2.82	3.49(4)	126
NZ-HZ1_14•••O_15#4	0.91	1.98	2.84(5)	158.6
NZ-HZ2_14•••O_15#5	0.91	2.65	3.37(7)	137.3
NZ-HZ2_14•••OXT_15#5	0.91	2.13	2.92(6)	144.8
N-H0_15•••O_14#1	0.88	1.95	2.78(3)	156.2
Symmetry transformations used to generate equivalent atoms: #1: x,y-1,z; #2: -x-½,y-½,-z-1; #3: x,y+1,z; #4: -x-½,y-½,-z; #5: -x-½,y+½,-z				

DF(I)NKF(I)				
D-H···A	**d(D-H) (Å)**	**d(H···A) (Å)**	**d(D···A) (Å)**	**<(DHA) (°)**
N-H0A_11•••OD1_11	0.91	2.6	3.134(9)	118.5
N-H0A_11•••OD1_11#1	0.91	2	2.803(8)	146.8
N-H0B_11•••OD1_11#2	0.91	1.92	2.762(9)	153.5
CA-HA_11•••O_11#3	1	2.34	3.111(9)	133.5
N-H0_12•••O_11#3	0.88	2.14	2.976(8)	158.2
CA-HA_12•••O_12#2	1	2.39	3.284(9)	148.2
N-H0_13•••O_12#2	0.88	2.05	2.901(8)	161.1
CA-HA_13•••O_13#3	1	2.5	3.373(9)	146.1
CB-HB1_13•••OD1_13#2	0.99	2.56	3.423(10)	145.5
ND2-HD2B_13•••OD1_13#2	0.88	2.05	2.883(9)	158.2
N-H0_14•••O_13#3	0.88	1.96	2.837(8)	172.2
CA-HA_14•••O_14#2	1	2.58	3.443(10)	144.1
CD-HD1_14•••IZ_11	0.99	3.3	3.946(13)	124.2
NZ-HZ1_14•••O_15#4	0.91	1.89	2.798(14)	175.2
NZ-HZ2_14•••OXT_15#5	0.91	2.04	2.892(15)	155.9
N-H0_15•••O_14#2	0.88	1.96	2.833(9)	173.2
Symmetry transformations used to generate equivalent atoms: #1: -x+3/2,y-½,-z+1; #2: x,y-1,z; #3: x,y+1,z; #4: -x+3/2,y-½,-z+2; #5: -x+3/2,y+½,-z+2				

References

[1S] Lausi A., Polentarutti M., Onesti S., Plaisier J. R., Busetto E., Bais G., Barba L., Cassetta A., Campi G., Lamba D., Pifferi A., Mande S. C., Sarma D. D., Sharma S. M., Paolucci G., The European Physical Journal Plus, 2015, 130(43), 1-8.

[2S] Kabsch W. XDS. Acta Crystallographica Section D, 2010, 66(2), 125–132.

[3S] Sheldrick G. M., SADABS, University of Göttingen, Germany, 2012.

[4S] Zwart, P. H., Banumathi, S., Dauter, M. & Dauter, Z., Acta Crystallographica Section D, 2004, 60, 1958-1963.

[5S] McGeehan J. E., Carpentier P., Royant A., Bourgeois D., Ravelli R. B. G., Journal of Synchrotron Radiation, 2007, 14, 99-108.

[6S] Sheldrick G. M., Acta Crystallographica Section A, 2015, 71, 3-8.

[7S] Sheldrick G. M., Acta Crystallographica Section C, 2015, 71, 3-8.

[8S] Emsley P., Lohkamp, B., Scott, W. G., Cowtan K., Acta Crystallographica Section D, 2010, 66(4), 486–501.

[9S] Parsons, S., Flack, H. D., Wagner, T., Acta Crystallographica Section B, 2013, 69(3), 249-259.

[10S] Farrugia, L., Journal of Applied Crystallography, 2012, 45(4), 849-854.

[11S] Macrae F., Bruno I. J., Chisholm J. A., Edgington P. R., McCabe P., Pidcock E., Rodriguez-Monge L., Taylor R., van de Streek J., Wood P. A., Journal of Applied Crystallography, 2008, 41, 466–470.

[12S] Schrodinger L., The PyMOL Molecular Graphics System. Schrodinger, LLC. http://www.pymol.org, 2015.

[13S] A.L.Spek, Acta Cryst. 2009, D65, 148-155.

[14S] Fischer, R. X., Tillmanns, E. Acta Cryst. 1988, C44, 775-776.