CHEMISTRY
A European Journal

Supporting Information

# Crystal Structure of the DFNKF Segment of Human Calcitonin Unveils Aromatic Interactions between Phenylalanines

Arianna Bertolani[+],[a] Andrea Pizzi[+],[a] Lisa Pirrie,[a] Lara Gazzera,[a] Giulia Morra,[b] Massimiliano Meli,[b] Giorgio Colombo,[b] Alessandro Genoni,[c, d] Gabriella Cavallo,[a] Giancarlo Terraneo,[a] and Pierangelo Metrangolo*[a, b, e]

chem_201604639_sm_miscellaneous_information.pdf

# Crystal structure of the DFNKF segment of human calcitonin unveils aromatic interactions between phenylalanines

Arianna Bertolani,[a,†] Andrea Pizzi,[a,†] Lisa Pirrie,[a] Lara Gazzera,[a] Giulia Morra,[b] Massimiliano Meli,[b] Giorgio Colombo,[b] Alessandro Genoni,[c,d] Gabriella Cavallo,[a] Giancarlo Terraneo[a] and Pierangelo Metrangolo [*a,b,e]

a) Laboratory of Nanostructured Fluorinated Materials (NFMLab), Department of Chemistry, Materials, and Chemical Engineering "Giulio Natta", Politecnico di Milano, Via L. Mancinelli 7, 20131 Milano, Italy.
b) Istituto di Chimica del Riconoscimento Molecolare, CNR, Via Mario Bianco 9, 20131 Milano, Italy.
c) CNRS, Laboratoire SRSMC, UMR 7565, Vandoeuvre-lès-Nancy, F-54506, France.
d) Université de Lorraine, Laboratoire SRSMC, UMR 7565, Vandoeuvre-lès-Nancy, F-54506, France.
e) HYBER Centre of Excellence, Department of Applied Physics, Aalto University, P.O. Box 15100, FI-02150, Espoo, Finland.

† These two authors contributed equally.

## Electronic Supplementary Information (ESI)

**Table of contents**

**S1. Materials**

The DFNK(*para*-iodophenylalanine) peptide without any modifications of the N and C termini was purchased from Biopeptek (Malvern, USA) with reported purity ≥ 98% and used without further purification. The integrity of the peptide was confirmed by ion spray mass spectrometry and the purity by reverse phase high-pressure liquid chromatography.

**S2. Computational analysis Molecular Dynamics (MD) - monomers in solution**

GROMACS 4.6.3 was used for the MD simulations.[1] DFNKF and DFNKF(I) peptides were modeled in extended conformations and inserted in a fully solvated cubic box with a volume of 81 $nm^3$, containing 2647 TIP4P water molecules.[2] The systems were subjected to energy minimization using steepest descent algorithm for 5000 steps. A single trajectory was calculated, with a total simulation time of 1.1 microsecond. The first 500 ps were discarded to account for equilibration and frames were saved every 2 ps. The calculation of the electrostatic forces utilized the PME implementation of the Ewald summation method. The LINCS[3] algorithm was used to constrain all bond lengths, while the SETTLE one for the water molecules.[4] A dielectric permittivity, $\epsilon = 1$, and a time step of 2 fs were chosen. An initial velocity, obtained from a Maxwellian distribution at the desired initial temperature of 300 K, were given to all atoms. The density of the system was adjusted performing the first equilibration runs at NPT condition by weak coupling to a bath of constant pressure ($P_0$ = 1 bar, coupling time $\tau_P$ = 0.5 ps).[5] In all simulations the temperature was maintained close to the intended values by weak coupling to an external temperature bath[6] with a coupling constant of 0.1 ps. OPLS-AAx force field was used to model the peptide which includes some modifications to allow the analysis of the σ-hole of the halogenated residues involved in the potential halogen bond.[7-9] The clustering was performed using the g_cluster tool in Gromacs.

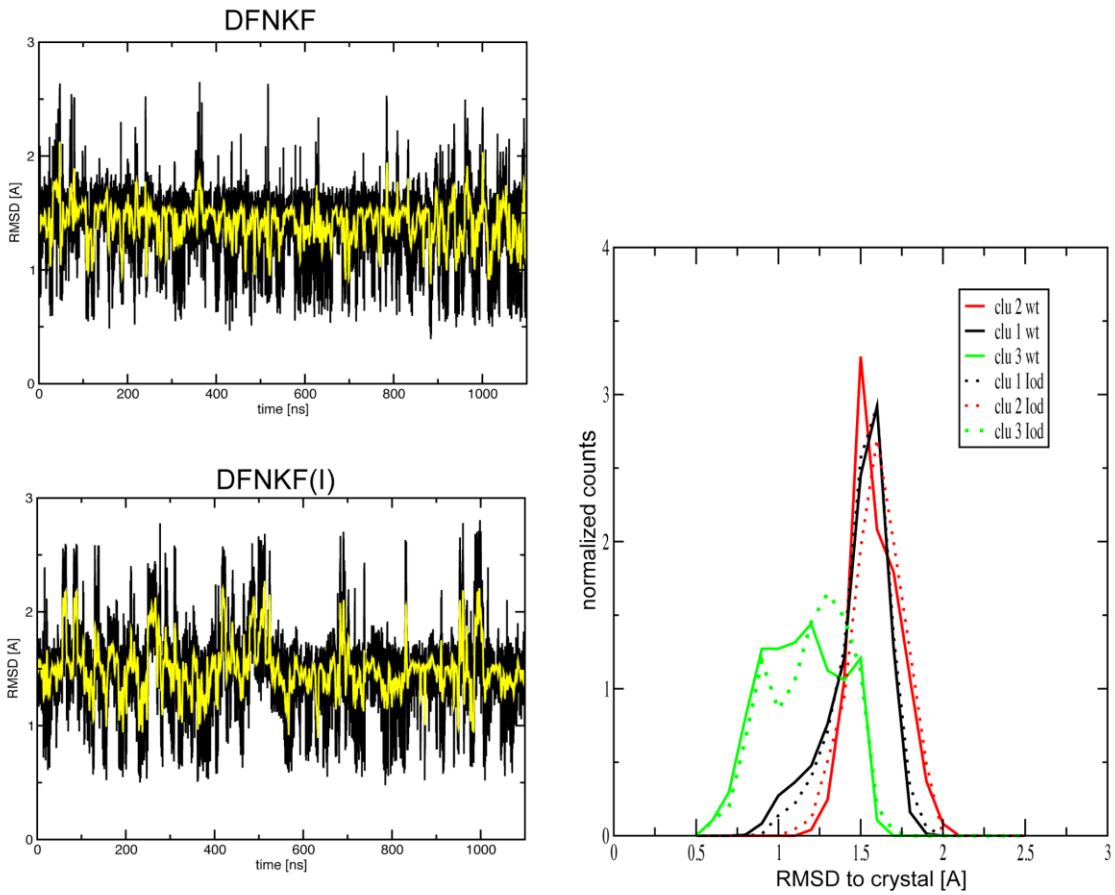



**Fig. S1** Left RMSD evolution calculated with respect to the backbone of the crystal structure. Right, histograms showing the distribution of RMSD values for members of the three most populated clusters. Solid lines, DFNKF. Dotted lines, DFNKF(I).

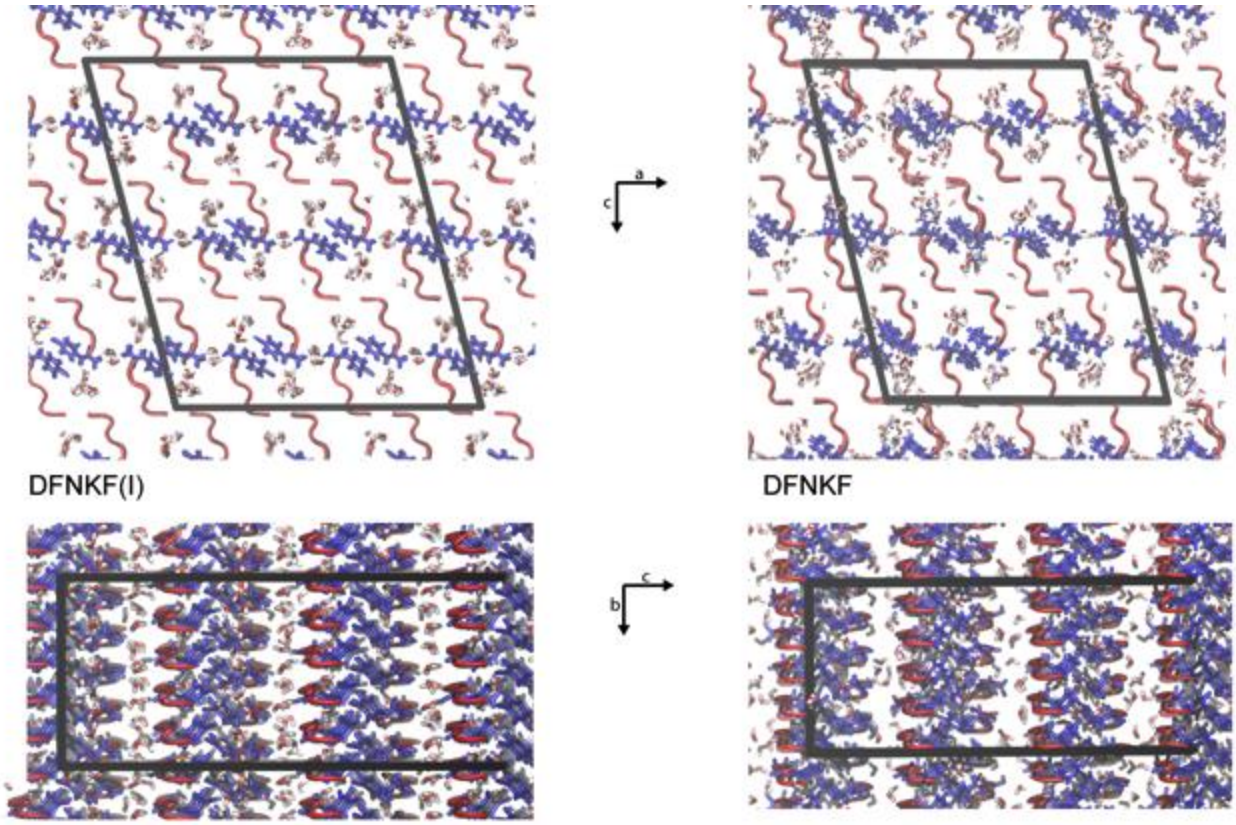



**Fig. S2** Final snapshot (t = 100 ns) of the crystal MD simulation of iodinated system (left) and of the model of wild type DFNKF system (right).

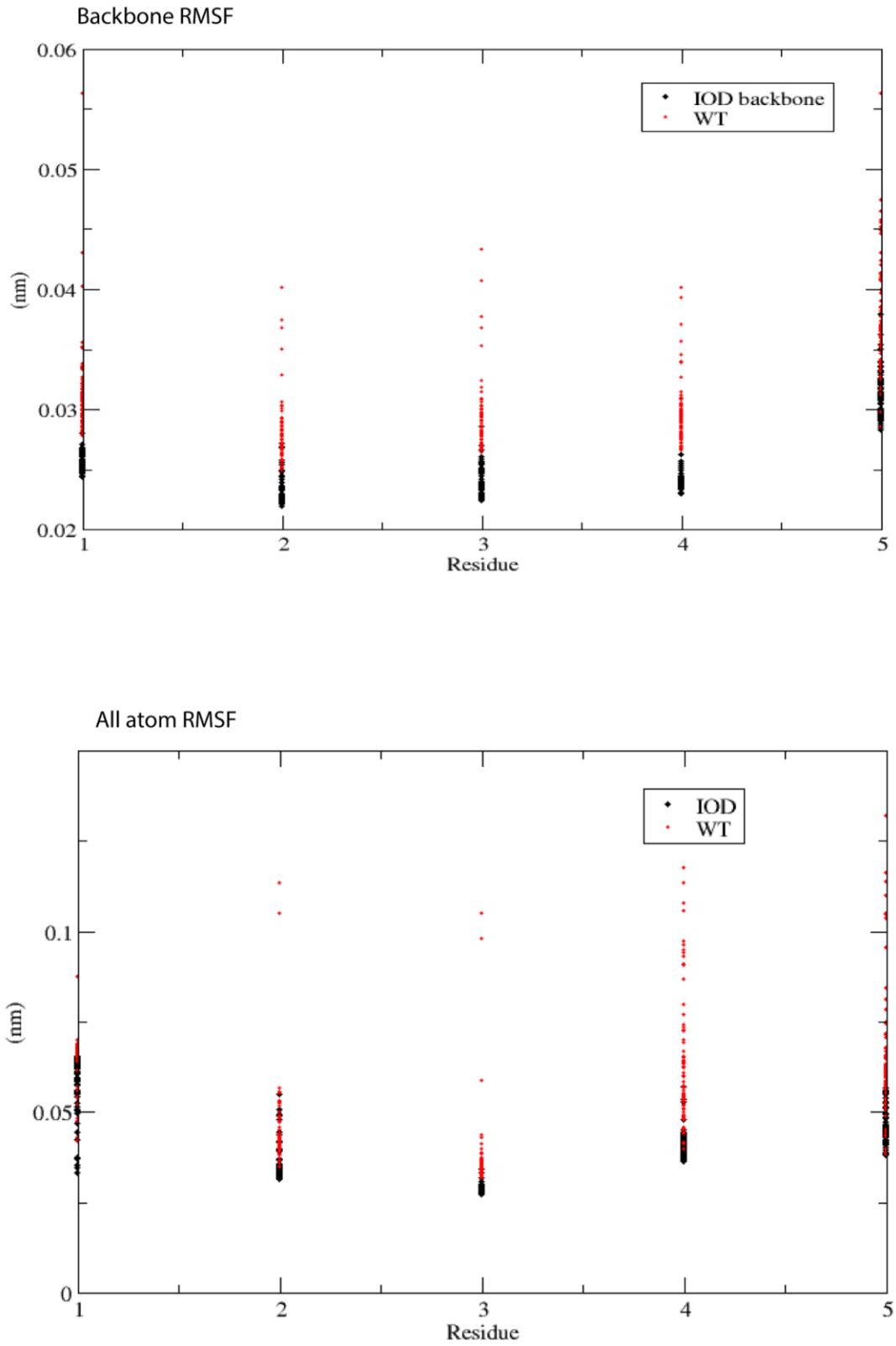



**Fig. S3** Root Mean Square Fluctuation (RMSF) calculated over the MD trajectory for each residue position along the crystal, DFNKF (red dots) and DFNKF(I) (black dots). Each dot refers to a single peptide.

**S3. Single-crystal X-ray diffraction analysis**

**S3.1 Peptide crystallization.** DFNK(*para*-iodophenylalanine, **DNFKF(I)**) crystals were grown using the sitting-drop vapor diffusion method. The peptide was dissolved in a 1:1 mixture of deionized water (18.2 MΩ · cm) and acetonitrile at room temperature at 1 mg · $ml^{-1}$. A drop of 20 μl has been deposited and put in vapor diffusion with a reservoir containing only deionized water (18.2 MΩ · cm). Single crystals were grown upon a six months period.

**S3.2 X-ray crystallography data acquisition.** Crystals of DFNK(*para*-iodophenylalanine) were extremely thin, curved yellow sheets or laminae. X-ray diffraction data were collected from DFNK(*para*-iodophenylalanine) crystals at Bruker APEX-II diffractometer equipped with micro-focus and CCD detector, using Mo-*K*α radiation (λ=0.71073 Å) and with a collection of 600”/frame. The crystal was cryo-cooled (103 K) for data collection (Bruker KRYOFLEX). The structure was solved by *SIR2002*[10] and refinements were carried out by full-matrix least-squares on $F^2$ using the *SHELXL*[11] program. The final set of data was comprised of 1460 unique reflections from a total of 2496 (with $I_o > 2\sigma(I_o)$ and $R_{ave} = 0.197$). Only iodine atom I1 was anisotropically refined, leading to a data/parameters ratio about 10 [6 considering only the data with $I_o > 2\sigma(I_o)$] from a total number of 225 parameters. To decrease the correlations in the refinement some restraints on parameters were imposed, namely the similarity of the displacement parameters in the whole molecule and the planarity of the two benzene groups. All the H atoms of the polypeptide were placed at calculated positions, while those of the water molecules (partially derived from the difference map and partially calculated on the basis of the shortest O···O contacts) were refined with strong geometric restraints. The main crystallographic data are reported in Table S1.

**Table S1.** Crystallographic data for the DFNKF(I) crystal.

| **Data collection and refinement** | **DFNK(*para*-iodophenylalanine)** |
|---|---|
| Chemical Formula | $C_{32}H_{42}IN_7O_9 \cdot 4H_2O$ |
| Temperature (K) | 103 |
| Radiation, wavelength (Å) | MoKα, 0.71073 |
| Space group | $P2_1$ |
| Crystal system | Monoclin |
| Resolution (Å) | 1.15 |
| Unit cell dimensions | |
| *a* (Å) | 18.639(9) |
| *b* (Å) | 4.832(2) |
| *c* (Å) | 21.953(11) |
| *β* (°) | 104.535(16) |
| Cell volume (Å$^3$) | 1914.0(16) |
| $d_{calc}$ (Mg m$^{-3}$) | 1.506 |
| *Z* | 2 |
| Crystal size (mm$^3$) | < 0.01 x 0.06 x 0.60 |
| $\mu$(Mo*K*α) cm$^{-1}$ | 0.908 |
| Data collected | 7995 |
| Unique data | 2496 |
| $R_{ave}$ | 0.197 |
| No. Obs. Data [$I_o > 2\sigma(I_o)$] | 1460 |
| No. parameters | 225 |
| No. restrains | 61 |
| $R_{all}$ | 0.166 |
| $R_{obs}$ | 0.104 |
| $wR_{all}$ | 0.243 |
| $wR_{obs}$ | 0.214 |
| Goodness-of-fit on $F^2$ | 1.015 |
| $\Delta\rho_{min,max}$ (e Å$^{-3}$) | -0.58, 1.04 |
| CCDC no. | 1504136 |

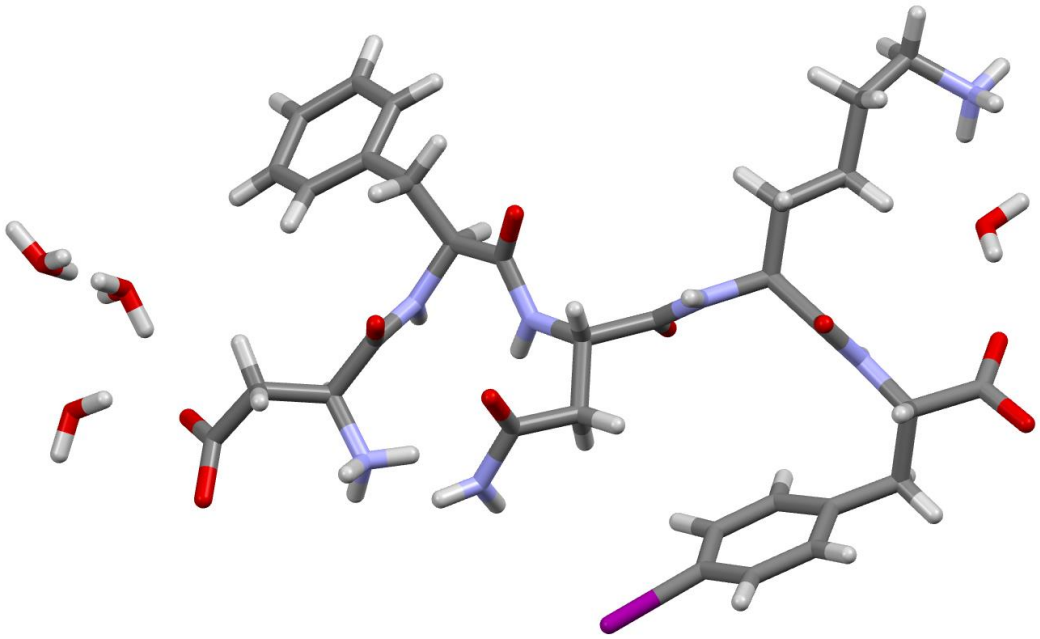


**Fig. S4** Asymmetric unit of DFNKF(I) consisting of a single peptide molecule solvated by four water molecules. Color code: C, grey; O, red; N, light blue; I, purple; H, white.

**Table S2.** Dihedral angles in the structure of DFNKF(I). Parallel β-sheets have average values of φ = - 119°, ψ = 113°.

| **Residue** | **Φ** | **ψ** | **Ω** | **$\chi^1$** | **$\chi^2$** | **$\chi^3$** | **$\chi^4$** |
|---|---|---|---|---|---|---|---|
| **Asp** | / | 106 | 173 | -72 | -82, 95 | / | / |
| **Phe** | 58 | 35 | 177 | -63 | -100, 82 | / | / |
| **Asn** | -89 | 146 | -177 | -63 | -56, 130 | / | / |
| **Lys** | -119 | 118 | -176 | -169 | 180 | -160 | -49 |
| ***p*-I-Phe** | -114 | / | / | -64 | -86, 101 | / | / |

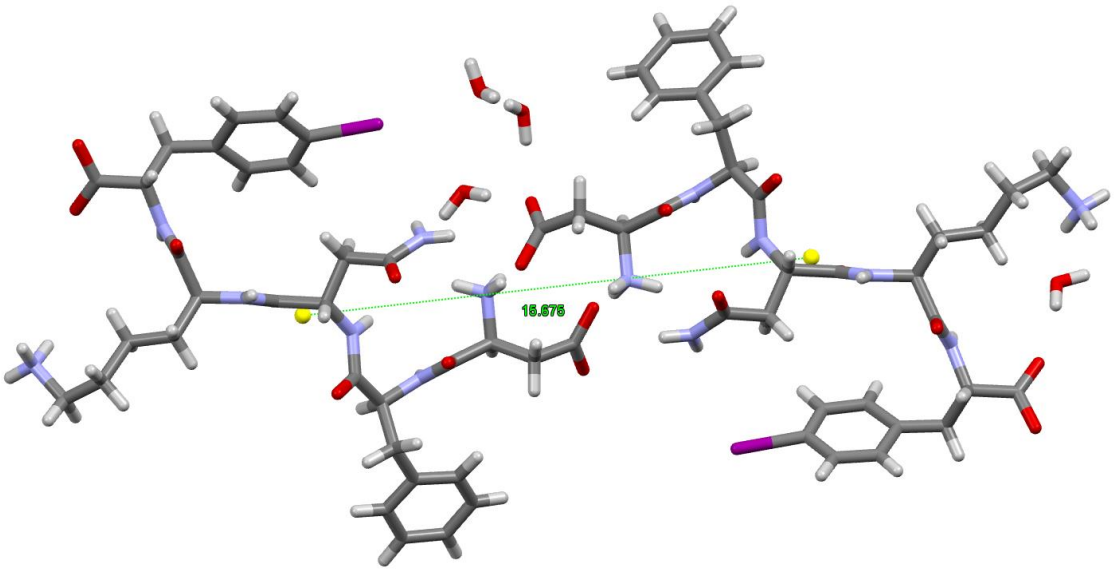



**Fig. S5** Crystal structure of DFNKF(I) view along the crystallographic b-axis. The wet interface is lined with water molecules that completely separates **DFNKF(I)** molecules, whereby the separation of sheets was around 15.7 Å (centroid-centroid distance). Colour code: C, grey; O, red; N, light blue; I, purple; H, white. HBs are in black dotted lines. The centroid of the molecule is depicted in yellow dot.

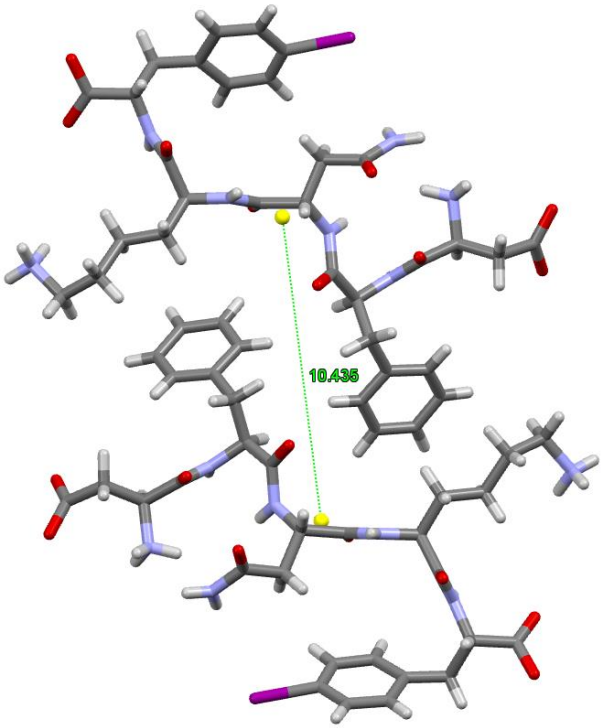



**Fig. S6** Crystal structure of DFNKF(I) view along the crystallographic b-axis. The dry interface contains no water molecules and these sheets were closer together, separated by around 10.4 Å (centroid-centroid distance). Colour code: C, grey; O, red; N, light blue; I, purple; H, white. HBs are in black dotted lines. The centroid of the molecule is depicted in yellow dot.

**S4. Non-Covalent Interaction (NCI) analysis on the DFNKF(I) crystal**

Non-Covalent Interaction (NCI) analyses[12-13] were performed on model-systems extracted from the crystal structure of the DFNKF(I) peptide (Figure S7) and the missing hydrogen atoms of the water molecules were added using the rules present in the hydrogen Gromacs database[14]. Afterwards, exploiting the Gaussian09 quantum chemistry package[15], for each system we carried out DFT calculations using the B3LYP functional[16] and the DZP basis-set[17]. These quantum mechanical computations provided us the electron distributions that were employed in the NCI calculations performed by means of the software NCIPLOT-3.0[18]. In particular, in preliminary NCI computations the following options were adopted: a) default cubic grid with a minimum distance of $\pm$ *2* Å from the outermost *x, y, z* coordinates, b) default increment of 0.1 atomic units (a.u.) along the *x, y, z* directions, c) cutoff interval for the electron density set to [-0.2 a.u., 0.2 a.u.] and d) cutoff for the reduced-gradient ***s(r)*** equal to 2.0 a.u. These calculations provided the characteristic two-dimensional plots *s(ρ)* in which the reduced-gradient ***s(r)*** is represented as a function of the electron density ***ρ(r)*** previously multiplied by the sign of second eigenvalue ($\lambda_2$) of the electron density Hessian matrix. The resulting graphs (Figures S8A-S8F) show the presence of non-covalent interactions (namely, peaks in the [-0.05 a.u., 0.05 a.u.] electron density range) for all the investigated systems. Therefore, further NCI computations were subsequently performed in order to better characterize the different interactions through the visualization of the corresponding reduced-gradient isosurfaces (Figures S9-S12). For these second-level NCI calculations, the same default cubic grid and increment of the preliminary computations were adopted, but a narrower electron density cutoff interval (i.e., [-0.05 a.u., 0.05 a.u.]) and a reduced-gradient cutoff set to 0.5 a.u. (isosurface value) were used.

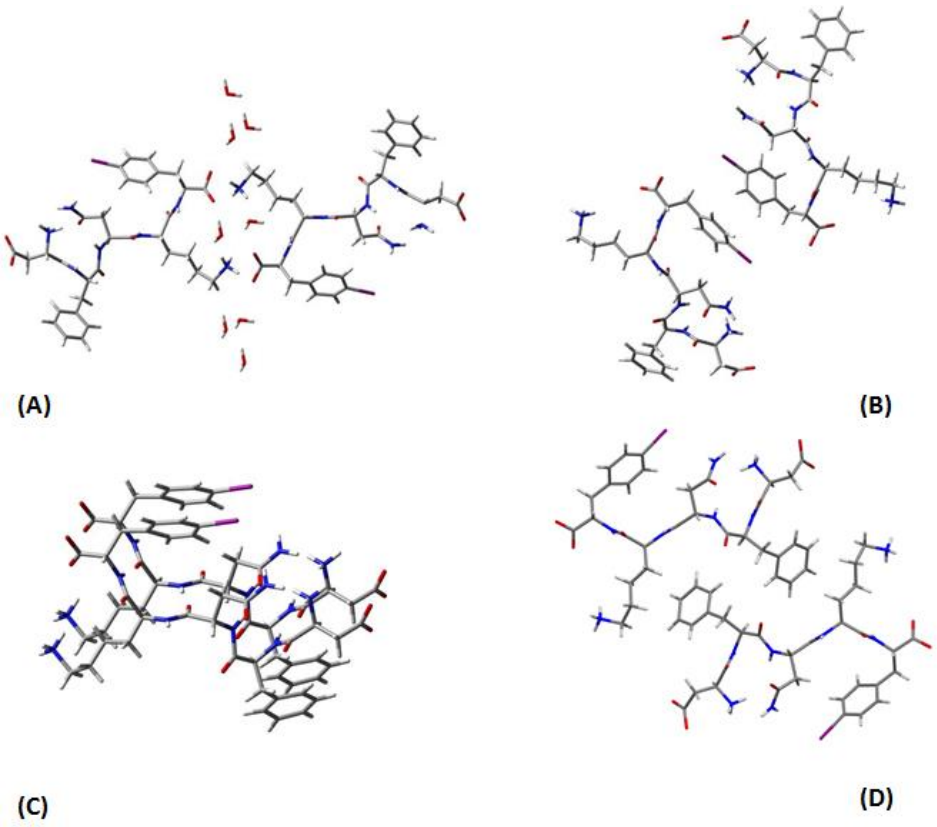



**Fig. S7** Iodinated model-systems extracted from the crystal of the DFNKF(I) peptide to perform NCI analyses: (A) A dimer of the unit-cell viewed down the crystallographic *a* axis; (B) Peptide dimer along the *a* crystallographic axis; (C) peptide dimer along the *b* crystallographic axis; (D) A dimer in the steric zipper.

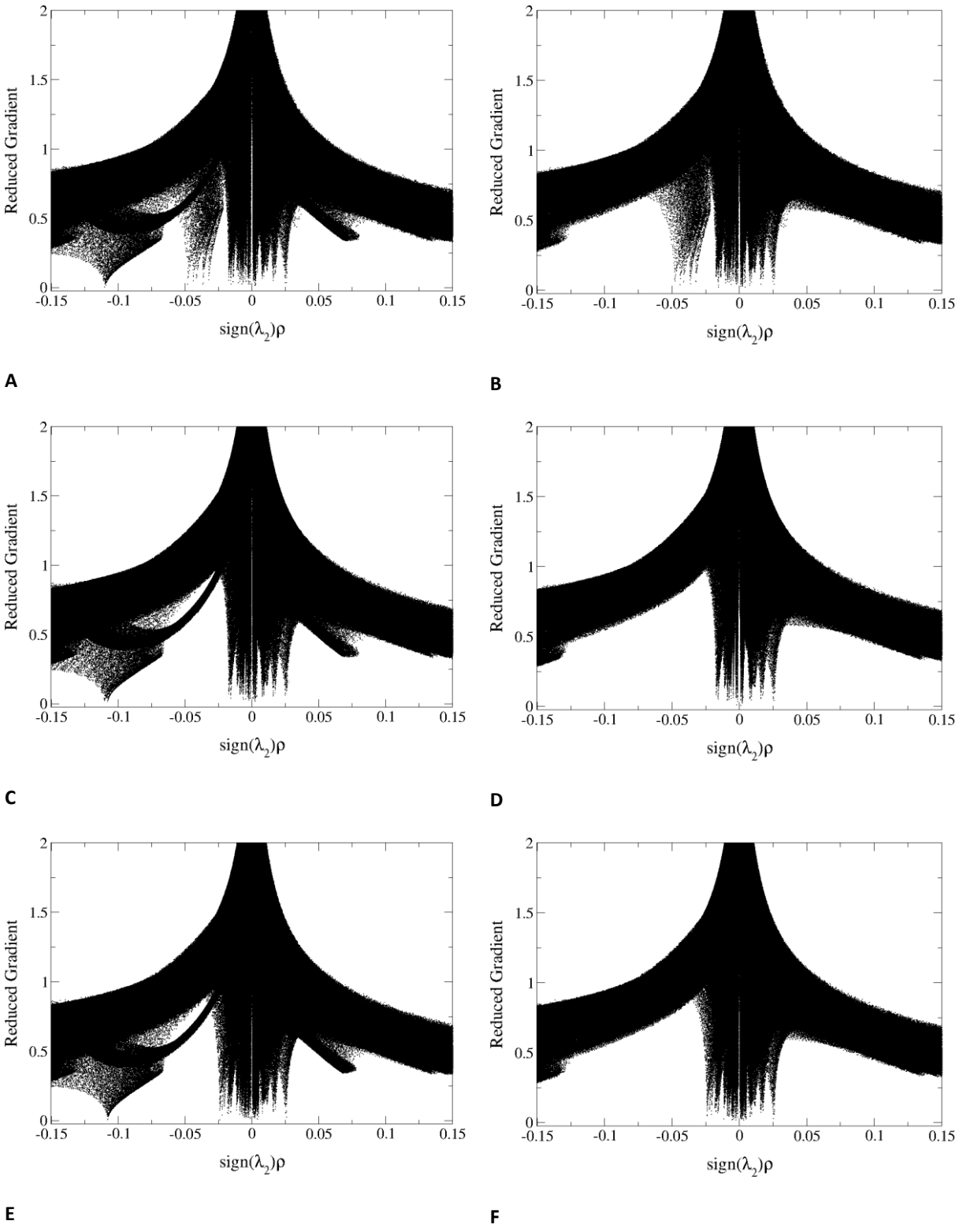
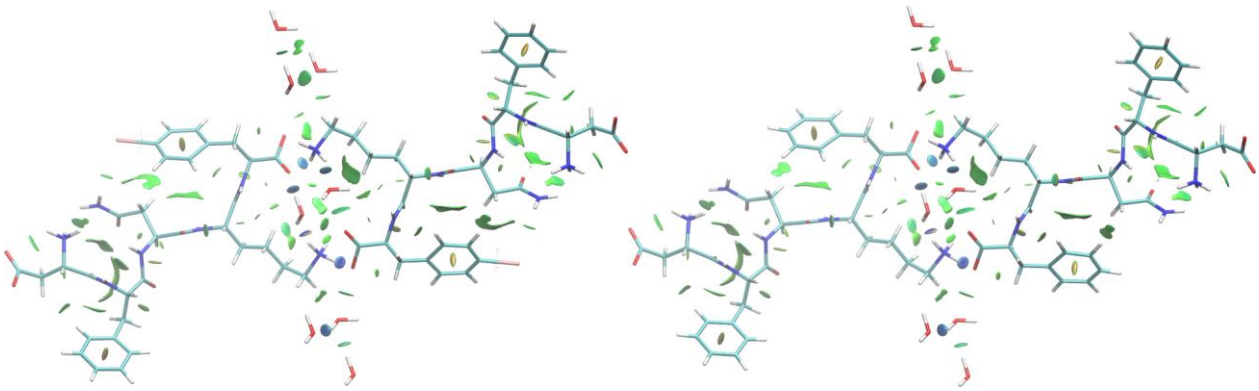
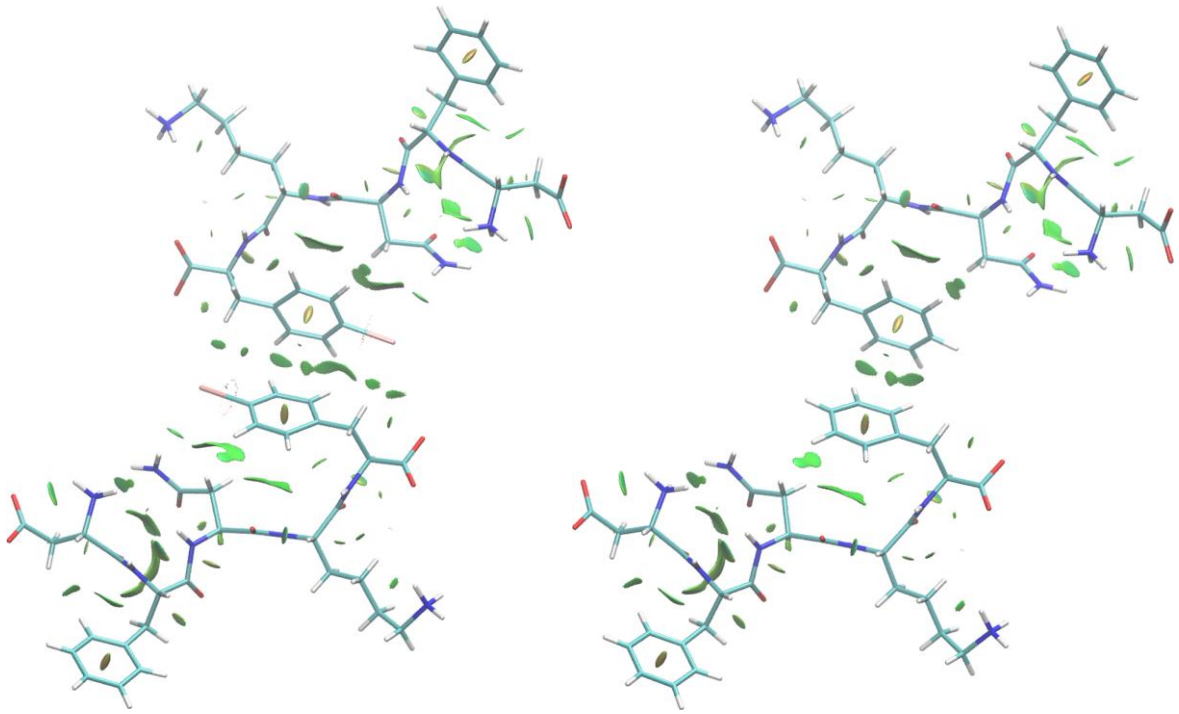



**Fig. S8** Characteristic two-dimensional plots $s(\rho)$ obtained at B3LYP/DZP level for (A) the iodinated unit-cell, (B) the non-iodinated unit-cell, (C) the iodinated peptide dimer along the *a* axis, (D) the non-iodinated peptide dimer along the *a* axis, (E) the iodinated peptide dimer along the *b* axis, (F) the non-iodinated peptide dimer along the *b* axis. Both the reduced gradient $s$ and $sign(\lambda_2)\rho$ are reported in au.



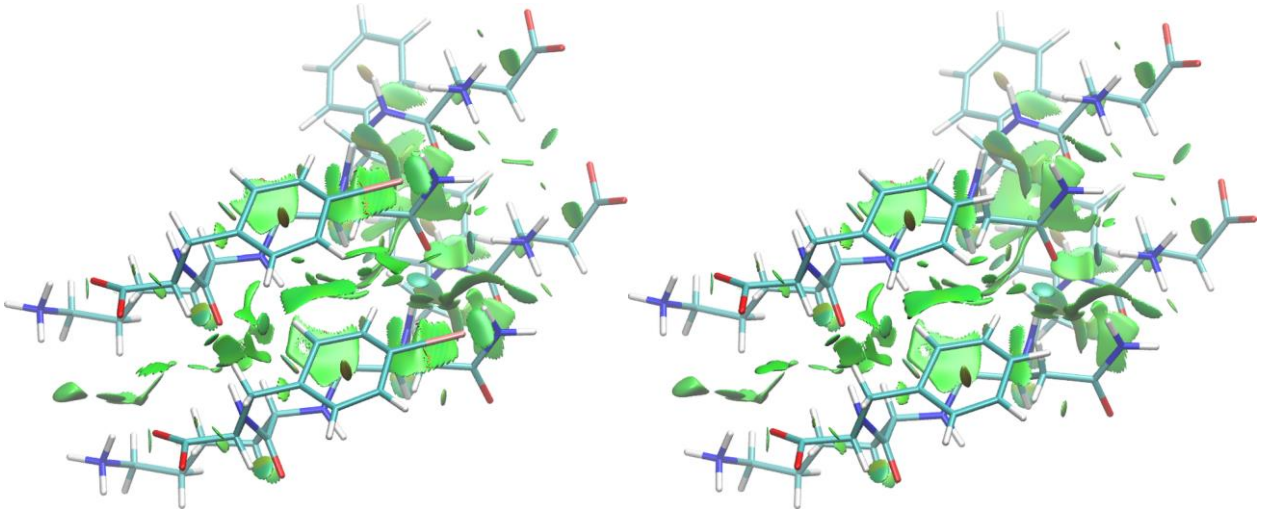
A B

**Fig. S9** Interaction (reduced-gradient) isosurfaces at the B3LYP/DZP level for (A) the iodinated unit-cell, (B) the non-iodinated unit-cell. The cutoffs are set at $s = 0.5\ au$ and $-0.05\ au < sign(\lambda_2)\rho < 0.05\ au$. The surfaces are colored on a blue-green-red scale according to the values of $sign(\lambda_2)\rho$: blue indicates strong attractive interactions, green indicates van der Waals interactions and red indicates strong non-bonded overlaps.



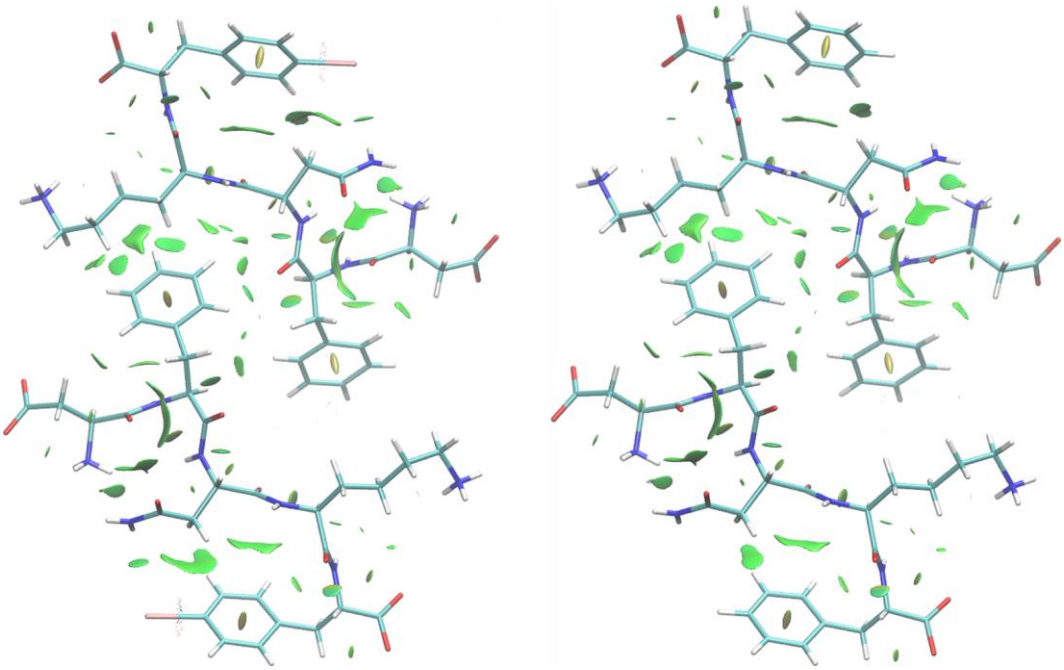
A B

**Fig. S10** Interaction (reduced-gradient) isosurfaces at the B3LYP/DZP level for (A) the iodinated peptide dimer along the *a* axis, (B) the non-iodinated peptide dimer along the *a* axis. The cutoffs are set at $s = 0.5\ au$ and $-0.05\ au < sign(\lambda_2)\rho < 0.05\ au$. The surfaces are colored on a blue-green-red scale according to the values of $sign(\lambda_2)\rho$: blue indicates strong attractive interactions, green indicates van der Waals interactions and red indicates strong non-bonded overlaps.



A B

**Fig. S11** Interaction (reduced-gradient) isosurfaces at the B3LYP/DZP level for (A) the iodinated peptide dimer along the *b* axis, (B) the non-iodinated peptide dimer along the *b* axis. The cutoffs are set at $s = 0.5\ au$ and $-0.05\ au < sign(\lambda_2)\rho < 0.05\ au$. The surfaces are colored on a blue-green-red scale according to the values of $sign(\lambda_2)\rho$: blue indicates strong attractive interactions, green indicates van der Waals interactions and red indicates strong non-bonded overlaps.



A B

**Fig. S12** Interaction (reduced-gradient) isosurfaces at the B3LYP/DZP level for (A) the steric zipper of the iodinated peptide, (B) the steric zipper of the non-iodinated peptide. The cutoffs are set at $s = 0.5\ au$ and $-0.05\ au < sign(\lambda_2)\rho < 0.05\ au$. The surfaces are colored on a blue-green-red scale according to the values of $sign(\lambda_2)\rho$: blue indicates strong attractive interactions, green indicates van der Waals interactions and red indicates strong non-bonded overlaps.

**S5. Computational analysis of the crystal model and molecular dynamics.**

The model of the crystal was created using the procedure described by Case *et al.*[19] The asymmetric units (ASU) described in supplementary Table 2 were replicated to obtain a "supercell" of 3x5x3 using the program Chimera (UCSF Chimera--a visualization system for exploratory research and analysis).[20] The obtained "supercell" contains 90 peptide fragments and 360 water molecules. The supercell was used to define the central tetrahedral simulation box, of size (55.917 Å, 24.160 Å, 65.589 Å), which was subjected to periodic boundary conditions (PBCs). Using this setup the virtual PBC images of the system were in contact with the main box, mimicking an infinite crystal. 88 TIP4P solvent molecules were added to the crystal waters in order to fill up the empty spaces using the GROMACS 4.6.3 tool genbox. The simulation was done using GROMACS 4.6.3 molecular simulation suite, the force field and the set up simulation was the same used for the monomer in solution. A hypothetical model of the crystal structure of the non-iodinated peptide was built by replacing the iodinated Phe with a standard Phe in the monomeric structure and then following the same procedure described above. After removing the crystal water molecules, the system was resolvated with a total number of 384 TIP4P water molecules using genbox. Both iodinated and non-iodinated crystal structures were simulated at 300 K using the same protocol used for the monomer in solution. The total simulation time for each system, after 5000 minimization steps using the SD algorithm and temperature equilibration, amounts to 100ns, corresponding to 25000 frames.

## S6. References


1. D. van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark and H. J. C. Berendsen, *J. Comp. Chem.*, 2005, **26**, 1701.
2. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926.
3. H. J. C. Berendsen, J. R. Grigera and T. P. Straatsma, *J. Phys. Chem.*, 1987, **91**, 6269.
4. S. Miyamoto and P. A. Kollman, *J. Comp. Chem.*, 1992, **13**, 952.
5. H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. Dinola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684.
6. W. L. Jorgensen, D. S. Maxwell and J. Tirado-Rives, *J. Am. Chem. Soc.*, 1996, **118**, 11225.
7. W. L. Jorgensen and P. Schyman, *J. Chem. Theory Comput.*, 2012, **8**, 3895.
8. P. Politzer, J. S. Murray and T. Clark, *Phys. Chem. Chem. Phys.*, 2013, **15**, 11178.
9. P. Auffinger, F. A. Hays, E. Westhof and P. S. Ho, *PNAS*, 2004, **101**, 16789.
10. M. C. Burla, M. Camalli, B. Carrozzini, G. L. Cascarano, C. Giacovazzo, G. Polidori and R. Spagna, *J. Appl. Cryst*., 2003, **36**, 1103.
11. G.M. Sheldrick, *J. Appl. Cryst*., 2011, **44**, 1281.
12. E. R. Johnson, S. Keinan, P. Mori-Sánchez, J. Contreras-García, A. J. Cohen and W. Yang, *J. Am. Chem. Soc.,* 2010, **132**, 6498.
13. J. Contreras-García, E. R. Johnson, S. Keinan, R. Chaudret, J.-P. Piquemal, D. N. Beratan and W. Yang, *J. Chem. Theory Comput.,* 2011, **7**, 625.
14. H. Bekker, H. J. C. Berendsen, E. J. Dijkstra, S. Achterop, R. van Drunen, D. van der Spoel, A. Sijbers, H. Keegstra, B. Reitsma, M. K. R. Renardus. Gromacs: a parallelcomputer for molecular dynamics simulations. In Physics Computing 92; Editors: R.A. de Groot and J. Nadrchal. World Scientific: Singapore, 1993.
15. Gaussian 09, Revision **D.01**, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2009.
16. a) A. D. Becke, *J. Chem. Phys.,* 1993, **98**, 5648; b) P. J. Stephens, F. J. Devlin, C. F. Chabalowski and M. J. Frisch, *J. Phys. Chem.,* 1994, **98,** 11623.
17. a) A. Canal Neto, E. P. Muniz, R. Centoducatte, F. E. Jorge, *J. Mol. Struct. (Theochem),* 2005, **718**, 219; b) C. L. Barros, P. J. P. de Oliveira, F. E. Jorge, A. Canal Neto, M. Campos, *Mol. Phys.,* 2010, **108**, 1965.
18. *NCIPLOT: A Program for Plotting Noncovalent Interaction Regions*. J.Contreras-García, E. R. Johnson, S. Keinan, R. Chaudret, J.-P. Piquemal, D. N. Beratan, and W. Yang. *J. Chem. Theory Comput.*, 2011 **7**, 625.
19. P. A. Janowski, D. S. Cerutti, J. Holton and D. A. Case, *J. Am. Chem. Soc.*, 2013, **135**, 7938.
20. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin *J Comput Chem.* 2004, **25**, 1605.