




Article

Hydrogen Bonds with Fluorine in Ligand–Protein Complexes—the PDB Analysis and Energy Calculations

Wojciech Pietruś , Rafał Kafel, Andrzej J. Bojarski  and Rafał Kurczab * 

Department of Medicinal Chemistry, Maj Institute of Pharmacology, Polish Academy of Sciences, Smetna 12, 31-343 Krakow, Poland; pietrus@if-pan.krakow.pl (W.P.); rafal.kafel@gmail.com (R.K.); bojarski@if-pan.krakow.pl (A.J.B.)

* Correspondence: kurczab@if-pan.krakow.pl; Tel.: +48-126-62-3301

Abstract: Fluorine is a common substituent in medicinal chemistry and is found in up to 50% of the most profitable drugs. In this study, a statistical analysis of the nature, geometry, and frequency of hydrogen bonds (HBs) formed between the aromatic and aliphatic C–F groups of small molecules and biological targets found in the Protein Data Bank (PDB) repository was presented. Interaction energies were calculated for those complexes using three different approaches. The obtained results indicated that the interaction energy of F-containing HBs is determined by the donor–acceptor distance and not by the angles. Moreover, no significant relationship between the energies of HBs with fluorine and the donor type was found, implying that fluorine is a weak HB acceptor for all types of HB donors. However, the statistical analysis of the PDB repository revealed that the most populated geometric parameters of HBs did not match the calculated energetic optima. In a nutshell, HBs containing fluorine are forced to form due to the stronger ligand–receptor neighboring interactions, which make fluorine the “donor’s last resort”.

Keywords: fluorine; PDB; hydrogen bonds; HBs



Citation: Pietruś, W.; Kafel, R.; Bojarski, A.J.; Kurczab, R. Hydrogen Bonds with Fluorine in Ligand–Protein Complexes—the PDB Analysis and Energy Calculations. *Molecules* **2022**, *27*, 1005. <https://doi.org/10.3390/molecules27031005>

Academic Editor: Miroslaw Jablonski

Received: 30 December 2021

Accepted: 29 January 2022

Published: 2 February 2022

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Fluorine is the most electronegative element, and this property has a significant impact on the bioavailability, lipophilicity, metabolic stability, acidity/basicity, and toxicity [1]. Since the second half of the 20th century [2], researchers have been exploring the possibility of using fluorinated molecules in medicine [1,3]. The important position of fluorinated molecules in medicinal chemistry can be understood by the exceptionally large number of fluorine-containing drugs currently available on the pharmaceutical market (Figure 1). The share of fluorinated compounds rose from 2% in 1970 to 8% in 1980, 13% in 1990, and reached 18% at the beginning of the 21st century. Among them, six products were in the “top-12” list and employed as anticancer, anti-inflammatory, analgesic, or antidepressant agents in medicine [4]. About 20% of the drugs used in 2010 contained fluorine atom(s) or fluoroalkyl group(s) [3], whereas in the last decade (2011–2020) 114 out of the 410 drugs approved by the Food and Drug Administration (FDA) (data from the Center for Drug Evaluation and Research (CDER)) [5] contained fluorine (Figure 1). Currently, fluorinated pharmaceuticals account for over 50% of the most profitable drugs (blockbuster drugs), and are also recognized as the best among the drugs used in almost all therapeutic areas [6].

The biological activity of drugs is determined by intermolecular interactions. These interactions also play an important role in stabilizing the ligand–biomolecule system. Hydrogen bonds (HBs), in particular, are considered to significantly influence the action of drug molecules on their targets [7–9]. Interestingly, fluorine or substituents containing this element have been shown to tune the intermolecular interactions in ligand–protein complexes [1,10]. Although characterized by high electronegativity, fluorine is a weak acceptor of HBs and, unlike other halogens, it is not a halogen bond (XB) donor in aromatic

systems [11,12]. However, the results of our previous study on small model systems (e.g., 2,6-difluoro-4-halogenoanilines) indicated that fluorine can act as a competitive and attractive acceptor for HBs and XBs as well as form F...F interactions [13]. Additionally, it is considered that fluorine-containing HBs are not typical and do not behave like conventional HBs (e.g., O...H-O and N...H-N), as demonstrated by a more angular nature and preference for less electronegative donors [14].

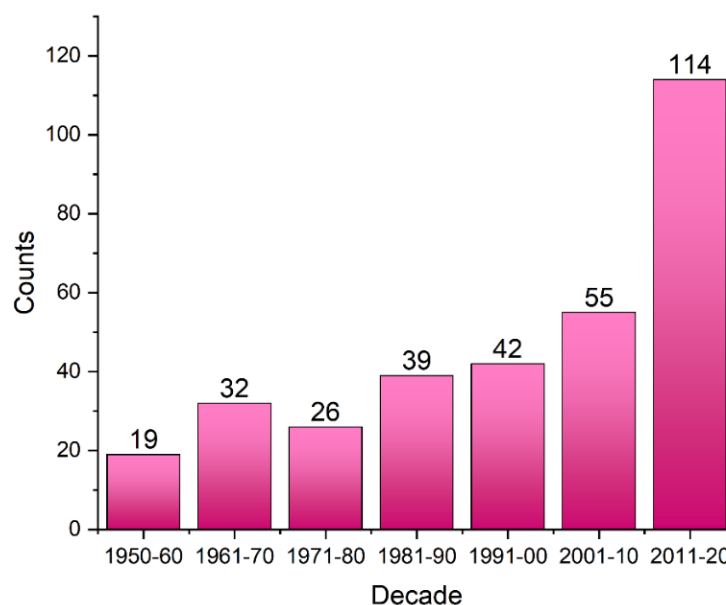


Figure 1. Number of marketed drugs containing fluorine per decade. Data were collected from the DrugCentral 2021 database (accessed 30 April 2021).

The biological activity of compounds can be tuned with the use of fluorine. However, there are no rules of thumb for predicting the preferred fluorine substitution sites in a molecule. Despite numerous studies on fluorine, the influence of this element on the pharmacodynamics properties of drugs remains unclear. A statistical analysis of the nature, geometry, and frequency of interactions occurring between fluorine in small molecules and the biological targets included in the Protein Data Bank (PDB) repository may allow understanding of the role of fluorine in ligand–receptor (L–R) complexes. Therefore, we carried out a wide statistical analysis and calculations to quantitatively and qualitatively explore the HBs (contacts) formed with fluorine in biological systems. The findings of this study may contribute to a thorough understanding of the effects of fluorine, to enable its rational use in drug design and for improving the efficiency of computational methods [15,16].

2. Results and Discussion

2.1. Choice of a Model System

We carefully chose the model systems used for database mining and quantum chemical calculations, with a focus on providing accurate representations of HBs with fluorine in the biological systems. Three types of HB donors were distinguished, namely hydroxyl, amine, and methyl group, and a pH of 7.4 ± 0.5 was considered to assess the protonation states of all entities. We extracted a ligand with interacting residue for performing high-level quantum chemical calculations in a reasonable time. This allowed us to determine the energy of isolated interaction with fluorine. For extracted amino acids from the main chain, the peptization reaction was reversed. Therefore, to maintain the proper structure of amino acids, missing atoms (hydroxyl to carboxylic group and hydrogen to nitrogen atom) were added and optimized with force field. Since 96% of the analyzed crystal structures were recorded with a resolution of $<3 \text{ \AA}$ (Figure 2), we analyzed all the collected structures.

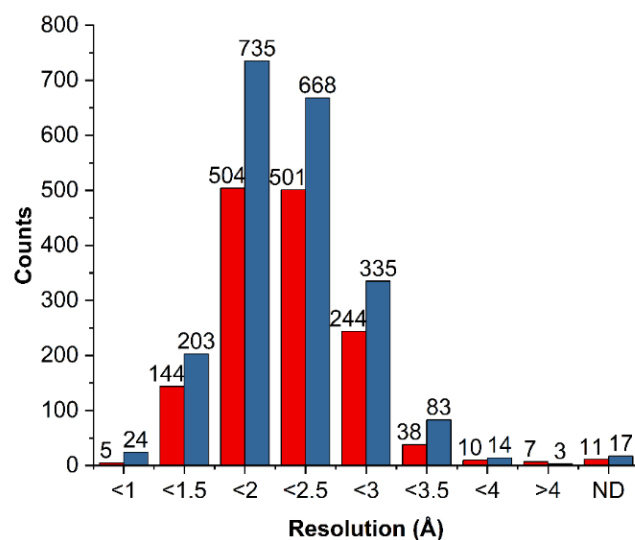


Figure 2. Statistical representation of the spectral resolution of the analyzed crystal structures deposited in the PDB database for L–R complexes with HB containing aliphatic fluorine (red bars) and aromatic fluorine (blue bars). “ND” (not determined) refers to the crystal structures obtained with the methods for which resolution was not specified (e.g., nuclear magnetic resonance).

2.2. General Statistics of HBs Containing Fluorine Atoms

It should be emphasized that the thresholds of HB geometric parameters considered in the statistical investigation based on the PDB data can significantly influence the results and conclusion. In this study, we assumed that the HB distance was $<4\text{Å}$ and the HB angle was $90^\circ\text{--}180^\circ$. Because we aimed to determine all contacts with fluorine atoms, these values can significantly exceed the standard geometric parameters of HBs; the distance can be below the sum of the van der Waals radii of interacting atoms, and the angle can differ by up to 120° . Complexes containing ligands with fluorine (from the LigandExpo repository) were extracted and analyzed. If the PDB entry contained more than one asymmetric unit (receptor oligomerization), the number of HBs with fluorine was multiplied by the number of occurrences of the same ligand. All measured HBs were used in further analysis, even if they came from the same PDB entry. The ligands were divided into two categories: molecules in which fluorine is bonded to an aliphatic carbon (F_{al}) and molecules containing fluorine bonded to an aromatic carbon (F_{ar}).

A total of 1787 (F_{al}) and 2324 (F_{ar}) unique PDB entries were found for fluorine-containing ligands. Based on the assigned boundaries and defined geometric thresholds, 165 interactions with a hydroxyl group, 612 with an amine group, and 3875 with a methyl group were identified for aliphatic fluorine (Figure 3A); and 121 interactions with a hydroxyl group, 606 with an amine group, and 6698 with H–C were identified for aromatic fluorine (Figure 3C). The number of $F\cdots\text{H-O}$ and $F\cdots\text{H-N}$ HBs was found to be larger for F_{al} , whereas the number of $F\cdots\text{H-C}$ HBs was two-fold higher for F_{ar} . For OH donors, three amino acids (SER, THR, TYR) were identified to be involved in HBs with F_{al} ; however, for F_{ar} , it appeared that TYR participates less frequently in HBs, which may be attributed to the greater acidity of the OH group.

The more significant differences were observed for NH donors, in which the amino acids commonly involved in HBs with F_{al} were in the order ARG>ASN>GLN>GLY>LYS (Figure 3B) and in HBs with F_{ar} were in the order ARG>GLY>LYS>GLN>ASN (Figure 3D). Surprisingly, glycine is the second most common amino acid, forming HBs with F_{ar} because the others are polar amino acids with a free amino group in their side chain.

By contrast, no clear preferences of amino acids in the occurrence of HBs with F_{al} and F_{ar} were observed in the case of CH donors. However, the results highlighted that fluorine most frequently forms HBs with nonpolar amino acids (LEU, VAL, PHE, ILE, ALA), implying that it prefers hydrophobic areas of binding pockets (Figure 3B,D).

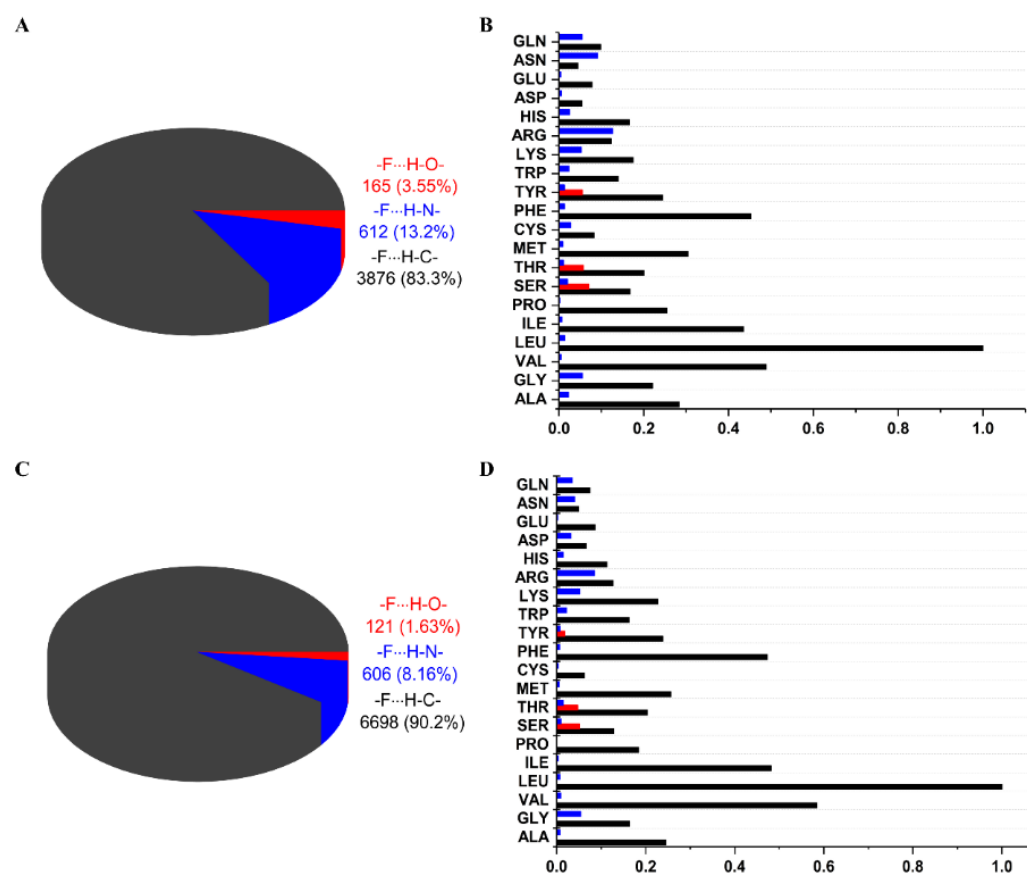


Figure 3. Number of hydrogen bonds found in the PDB repository to meet the boundaries for geometric parameters: (A) between F_{al} and OH, NH, and CH donors; (B) normalized division into interacting amino acids; (C) between F_{ar} and OH, NH, and CH donors; and (D) normalized division into interacting amino acids.

Based on this classification, we generated the density maps showing the geometric parameters of HBs (Figure 4). As only a small number of $F \cdots H-O$ HBs were identified, certain conclusions could not be drawn (Figure 4). The density maps of HB geometric parameters obtained for the NH and CH donors (as well as OH) revealed that fluorine prefers geometries with a distance of $>3 \text{ \AA}$ and an angle of 100° – 140° (Figure 4). However, it should be noted that more HBs were found for NH than OH donors, with a more linear geometry and short distances, but in many cases, those interactions are forced by the neighboring functional groups of a ligand interacting with amino acids.

In summary, fluorine-containing HBs reveal more angular geometric preferences than typical HBs (rather linear HBs $O \cdots H-O$, $N \cdots H-N$). Thus, in the next step, we explored the relationship between the geometry of $F \cdots H-X$ ($X = O, N, C$) bond and the energy contribution to the ligand–receptor complexes.

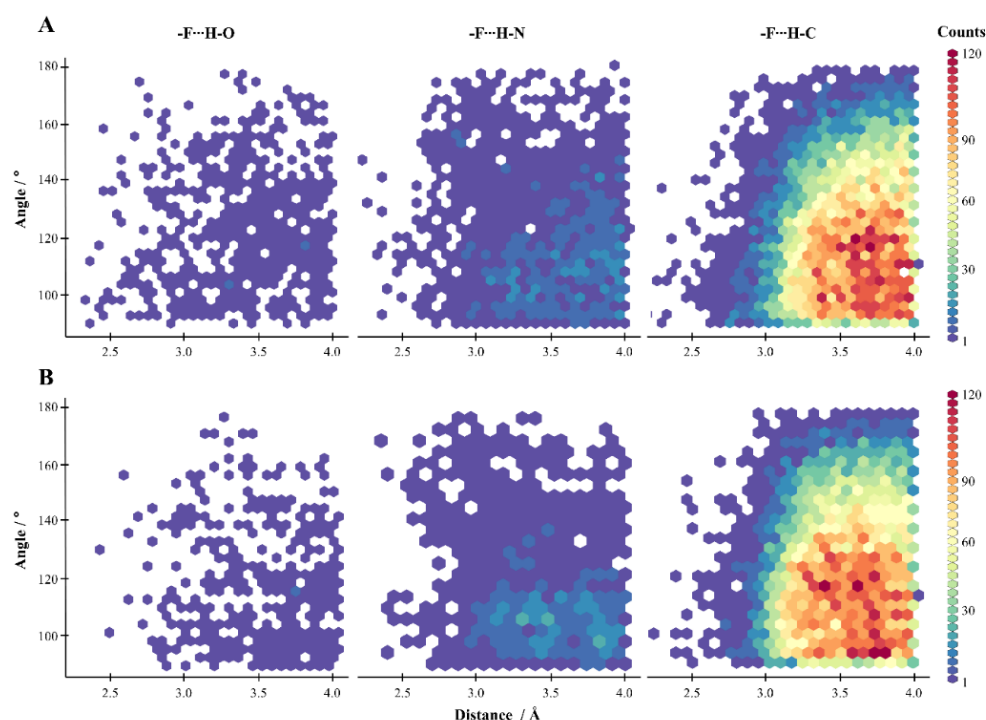


Figure 4. Density maps of geometric parameters of HBs with (A) F_{al} and (B) F_{ar} .

2.3. Energy of HBs with Fluorine

Ligand–receptor complexes are stabilized by various intermolecular forces, such as strong HBs ($O\cdots H-O$, $N\cdots H-O$, $N\cdots H-N$), weak HBs ($O\cdots H-C$, $S\cdots H-N$), XBs, π -stacking, salt bridge, amide stacking, cation- π , and hydrophobic interactions and others [17]. Fluorine-containing HBs, especially those with hydroxyl and amine donors, are not common in biological systems (Figures 3 and 4). Therefore, it is important to determine the strength and geometric preferences of these HBs in biological systems. In this study, we attempted to evaluate the nature and energetic dependencies of HBs with fluorine in the theoretical background by performing quantum chemical calculations using small molecular systems extracted from ligand–biomolecule crystals ($O-H$, $N-H$, $+N-H$, and $C-H$ were only considered to be HB donors). We determined the energy of HBs with fluorine found in crystal structures by applying three different methods as follows: (1) Diff—energy was calculated as the difference between the energy of the interacting molecules and the sum of the energies of isolated species calculated in Gaussian; (2) QTAIM—energy was calculated at BCP in AIMAll software; and (3) ETS—energy was calculated between two interacting molecules using the ETS-NOCV scheme implemented in ADF software.

At first, a simple statistical analysis was performed on the data obtained from the three approaches using the correlation coefficient and Pearson test in R (Supplementary Tables S1 and S2). The results of the analysis indicated the highest correlation between Diff and ETS methods ($p < 0.05$, correlation coefficient ~ 1) because they consider the energy of the entire system and approximately 70% of the calculated energy accounts for the same nature of interaction (attractive/repulsive). Additionally, the correlation decreased for stronger interactions (Supplementary Tables S1 and S2) due to the fact that the Diff method is intended for weak and medium HBs; strong HBs result in geometry distortion of the interacting molecules, which decreases the accuracy of the evaluation of the HB itself [18]. The energies of HBs calculated by QTAIM did not correlate with those determined by the remaining methods, since this method takes only isolated L–R interaction and neglects long-range interactions occurring between the atoms from separated fragments.

The distribution of calculated interaction energies for all selected complexes for a given type of HB with fluorine and method is illustrated in Figure 5. For HBs with F_{al} , the interaction energy calculated by QTAIM varied between 0 and -1.2 kcal/mol (the

weakest HBs were found for $F_{al} \cdots H-O$, with an energy value of -0.64 kcal/mol). For $F_{al} \cdots H-C$ HBs, the energy was $(-0.69$ kcal/mol). The strongest HBs (~ -0.8 kcal/mol) were observed for $F_{al} \cdots H-N$ (no significant difference was noted between $F_{al} \cdots H-N$ HB and $F_{ar} \cdots H-N^+$ HB) (Figure 5A). The energy range determined by the Diff method was also between 0 and -1.2 kcal/mol, while the energy determined by the ETS method ranged from 0 to -8 kcal/mol (Figure 5A). Additionally, the Diff method indicated that the charge-assisted $F_{al} \cdots H-N^+$ was the weakest HB (-0.70 kcal/mol), while $F_{al} \cdots H-N$ HB was stronger (-0.96 kcal/mol). Unlike Diff, the results of the ETS method showed that the charge-assisted $F_{al} \cdots H-N^+$ HB was the stronger (-5.11 kcal/mol), while $F_{al} \cdots H-N$ HBs were weaker than $F_{al} \cdots H-O$ HBs (-2.28 and -2.46 kcal/mol, respectively) (Figure 5A). For HBs with F_{ar} , a different trend was noted in the QTAIM method than for HBs with F_{al} , where the strongest HBs had OH as a donor (-0.94 kcal/mol). The $F_{ar} \cdots H-N$ HBs were found to be slightly weaker with median energy values of ~ -0.7 kcal/mol (for ^+NH and NH donors), and the weakest was $F_{ar} \cdots H-C$ HB (-0.62 kcal/mol) (Figure 5B). A comparison of Diff and ETS methods revealed a similar trend as in the case of F_{al} —the Diff method indicated that the $F_{ar} \cdots H-N^+$ HB was the weakest (-0.8 kcal/mol), while the ETS method indicated it as the strongest interaction (-5.3 kcal/mol) (Figure 5B). The QTAIM method showed that $F \cdots H-O$ HBs with F_{ar} were stronger than those with F_{al} , while HBs with other donors were found at a similar energy level. However, it should be emphasized that both Diff and ETS methods revealed higher stabilization energy for HBs with F_{ar} compared to HBs with F_{al} .

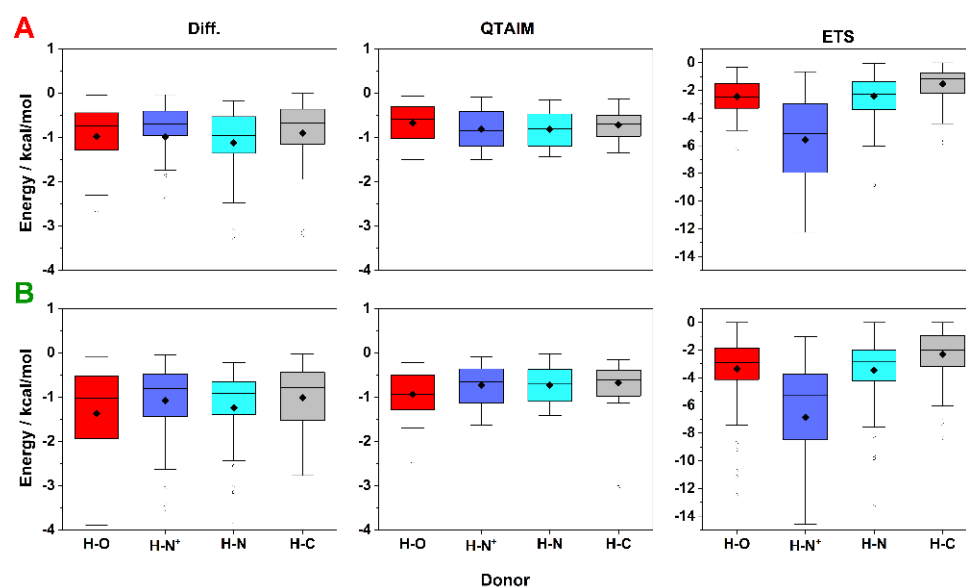


Figure 5. Box plots showing the distribution of the stabilization energy for HBs containing fluorine bonded to an (A) aliphatic or (B) aromatic carbon. A comparison is made for the individual donor groups (OH, ^+NH , NH, and CH) as well as the calculation approaches used (Diff, QTAIM, and ETS).

The hydroxyl donor occurs in the side chains of three amino acid—tyrosine (TYR), threonine (THR), and serine (SER). The phenolic hydroxyl group (TYR) is significantly more acidic (pK_a of about 9.8 in polypeptides) than the aliphatic hydroxyl group (SER or THR, $pK_a \sim 13.6$) [19]. In addition, Graton et al. found in an analysis of the PDB repository that the distances and angle of HBs with a hydroxyl group decreased in the order $THR > SER > TYR$, which suggests that TYR forms the stronger HBs [20]. In the present study, the results obtained by the QTAIM method revealed that for F_{al} (Figure 6A), the energy of HBs does not depend on the $F \cdots H-O$ angle, as the highest values were observed in the whole range of the analyzed angles. Instead, the energy of $F_{al} \cdots H-O$ HBs closely correlated with the distance, as observed in the case of conventional hydrogen bonding. It should be mentioned that the QTAIM method showed higher energy for $F \cdots H-O$ HBs

with aromatic fluorine than for HBs with aliphatic fluorine (Figure 6). Most of the $F \cdots H-O$ HBs with distances shorter than 2.75 Å are repulsive (Figure 6), which shows that despite the stabilizing nature of the $F \cdots H-O$ HB itself, the interacting fragments have a positive energy contribution (repulsive character). This effect may be due to the interaction of the neighboring atoms, or high positive kinetic energy. All three methods showed that the highest stabilizing energies (red squares in Figure 6) were in the range of 2.85–3.45 Å (for both fluorine) and 150° – 120° for F_{al} and 145° – 120° for F_{ar} , suggesting that these are the optimal ranges of geometric parameters for $F \cdots H-O$ HBs. The analysis of the selected crystal structures did not show any significant differences between the $F \cdots H-O$ HBs of attractive and repulsive nature. The only differentiating factor identified was the $F \cdots O$ distance (Figure 6).

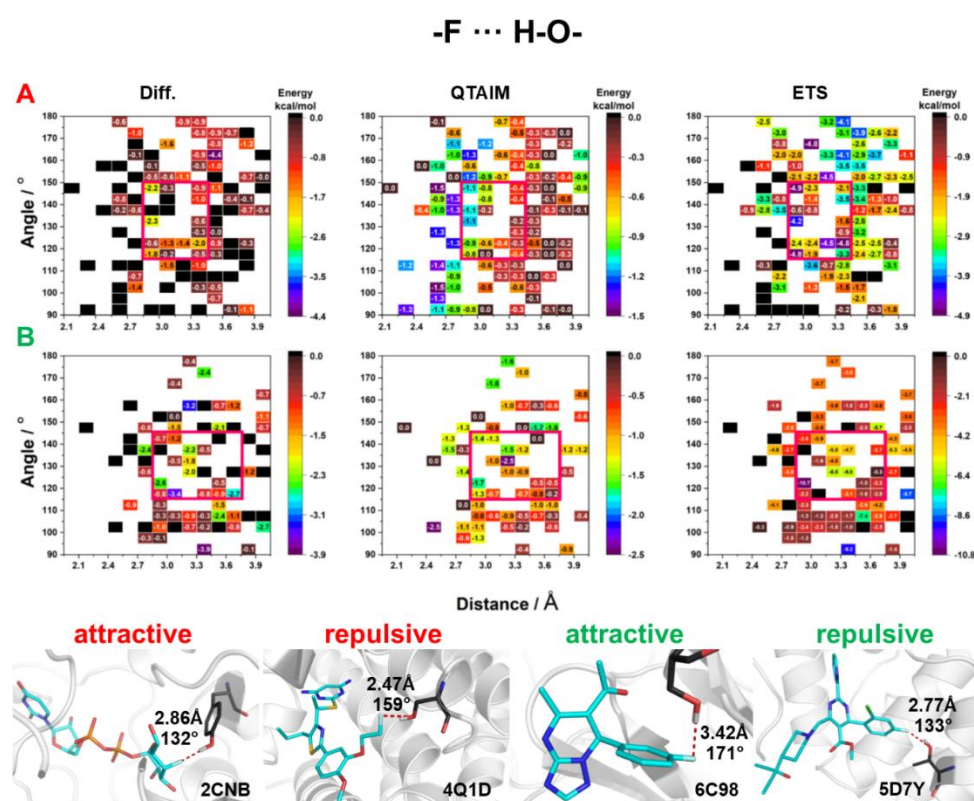


Figure 6. HB energy maps generated based on Diff, QTAIM, and ETS calculations of interaction energy between OH donor and fluorine attached to (A) an aliphatic fragment and (B) an aromatic ring for specific geometric parameters. The areas for which the highest stabilizing energy was observed in all three methods are marked with a red square.

Among amino acids, three (ARG, LYS, HIS) have additional amino groups. The side chain of arginine (ARG) is amphipathic because at physiological pH it contains a positively charged guanidine group ($pK_a = 12.48$). Another amphipathic amino acid is lysine (LYS), the side chain of which contains a positively charged primary amine group at the end of the long hydrophobic carbon tail ($pK_a = 10.53$). Histidine (HIS) contains an imidazole side chain. His pK_a is 6, above which one of the two protons is missing (in physiological pH, histidine has two tautomers). Since it is difficult to automatically protonate the appropriate nitrogen atom of histidine which forms an $F \cdots H-N^+$ HB, we calculated the energy of $F \cdots H-N^+$ HB only for LYS and ARG (Figure 7). The QTAIM calculations showed that the energy of $F_{al} \cdots H-N^+$ and $F_{al} \cdots H-O$ HBs was similar. Additionally, $F \cdots H-N^+$ was found to be strongly influenced by the distance between $F \cdots N$ and not by the angle (Figure 7) (as noticed for $F \cdots H-O$ HBs). A similar trend for both aliphatic and aromatic F was observed for $F \cdots H-N^+$ HBs, but the highest interaction energy was mostly localized at higher values of the $F_{ar} \cdots H-N^+$ angle (Figure 7). For aliphatic fluorine, the range of geometric parameters

in which the three methods indicated the strongest $F_{al} \cdots H-N^+$ HBs was 140° – 120° and 2.85–3.45 Å (as for $F \cdots H-O$ HB). In the case of F_{ar} , two ranges of geometric parameters were distinguished: (1) 170° – 150° and 3.0–3.6 Å; and (2) $< 120^\circ$ and 2.4–3.6 Å. Due to its large volume, the guanidine group interacts not only with fluorine directly but also with neighboring atoms. Therefore, the energy calculated by Diff and ETS methods might be overestimated.

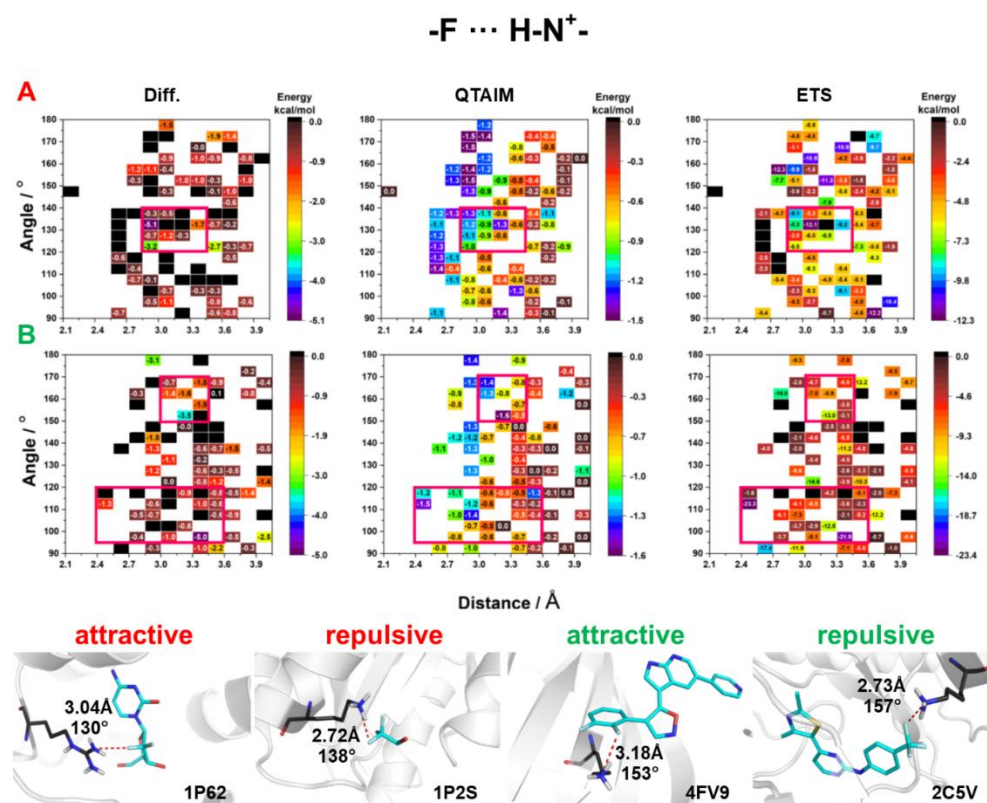


Figure 7. HB energy maps generated based on Diff, QTAIM, and ETS calculations of interaction energy between positively charged NH donor and fluorine attached to (A) an aliphatic fragment and (B) an aromatic ring for specific geometric parameters. The areas for which the highest stabilizing energy was observed in all three methods are marked with a red square.

Interestingly, the analysis of crystal structures with $F \cdots H-N^+$ HBs showed that for $F \cdots N^+$ distances of < 2.8 Å, almost 70% of HBs exhibited a destabilizing character. Moreover, the interaction of positively charged nitrogen with the CF_3 group, with a partial positive charge on the carbon atom, is often repulsive (Figure 7).

The $F \cdots H-N$ HBs were found almost four times more frequently in PDB than $F \cdots H-O$ HBs (Figures 3 and 4). The energies of $F \cdots H-N$ HBs were calculated for all amino acids, except for arginine and lysine as these amino acids contain positively charged nitrogen atoms. Furthermore, whether the nitrogen atom was in the main chain or the side chains (ASN, GLN) was not considered in the analysis. The energies of $F \cdots H-N$ HBs determined by the QTAIM method showed no significant differences between F_{al} and F_{ar} . In addition, it must be noted that energy is inversely proportional to the HB distance and does not depend on the $F \cdots H-N$ angle (Figure 8). However, since the nitrogen atom was mostly present in the main chain, and was thus adjacent to different atoms, the energies calculated by Diff and ETS methods mostly had a destabilizing nature, which might be due to steric effects. On the other hand, for F_{al} , the areas where the energies were found to be high and exhibited a stabilizing character had a narrow range of 150° – 125° and 2.4–3.45 Å, while for F_{ar} the areas were within the range of geometric parameters (165° – 135° and 2.85–3.75 Å) (Figure 8). The analysis of selected crystal structures showed no significant differences between the systems with attractive energy and repulsive energy (Figure 8).

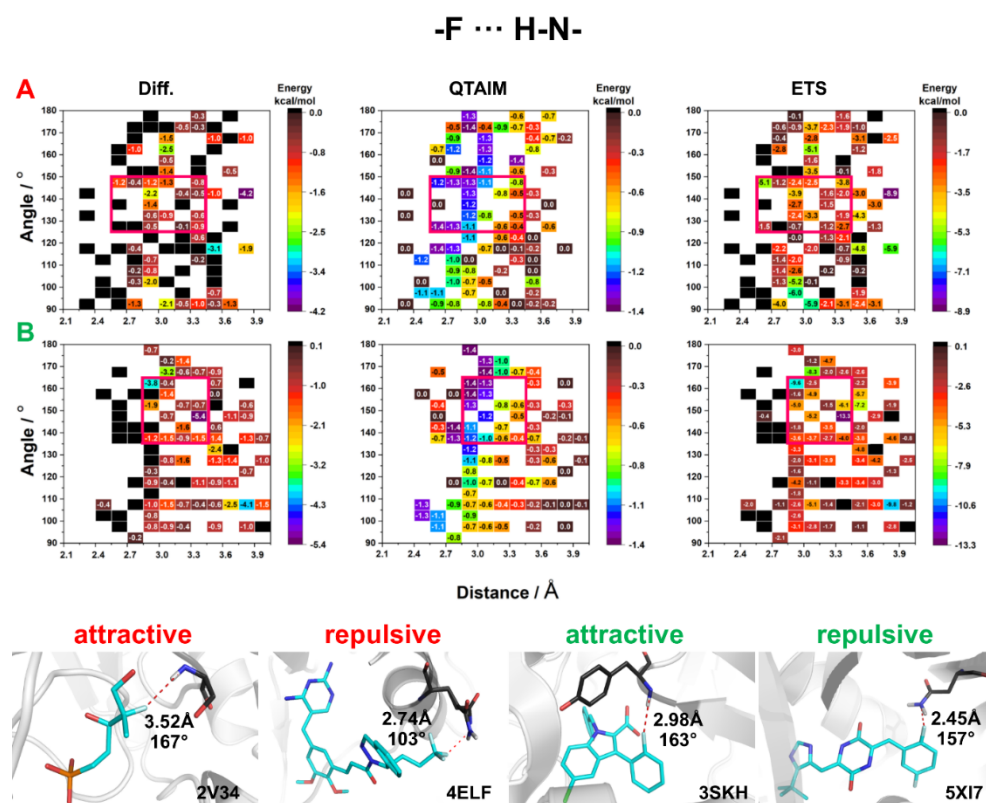


Figure 8. HB energy maps generated based on Diff, QTAIM, and ETS calculations of interaction energy between NH donor and fluorine attached to (A) an aliphatic fragment and (B) an aromatic ring for specific geometric parameters. The areas for which the highest stabilizing energy was observed in all three methods are marked with a red square.

The $F \cdots H-C$ HBs were found to be the most abundant in biological systems (Figures 3 and 4). The interaction energies calculated by the QTAIM method showed that $F_{al} \cdots H-C$ HBs are stronger than $F_{ar} \cdots H-C$ HBs. In addition, the interactions with an HB distance of $<2.7 \text{ \AA}$ showed a destabilizing character (Figure 9). The results produced by Diff and ETS methods were quite divergent, and it is difficult to find any constant trend. Interestingly, the results obtained from all three methods indicate that HBs with F_{al} are stronger than those with F_{ar} (Figure 9). The energetically favorable HBs with F_{al} had an angle of 155° – 145° and a distance of $>2.7 \text{ \AA}$, while HBs with F_{ar} had an angle of 160° – 120° and a distance of $>2.85 \text{ \AA}$ (red squares in Figure 9). The analysis of selected crystal structures showed that $F \cdots H-C$ HBs mostly exhibited a stabilizing character for distances longer than 3 \AA (Figure 9).

Analysis of the ETS-NOCV decomposition results showed that for uncharged donors (OH, NH, CH) the contribution of Coulomb energy term has the greatest impact on the stabilization energy of HBs with fluorine, while for NH^+ the XC energy term has the largest contribution. The reason for the destabilizing nature of the shorter HBs with fluorine may be due to the high value of the kinetic energy contribution (Figure S2). To determine the significance of HBs with fluorine, the density maps of geometric parameters of HBs found in the PDB repository (Figures 3 and 4) were compared with the corresponding HB energy maps (Figures 6–9). The proposed areas (red squares) of favorable geometric parameters with the highest energy values did not match with the most occupied areas of geometric parameters. This suggests that HBs with fluorine do not play a significant role in the stabilization of the L–R system and are often formed under unfavorable geometric parameters.

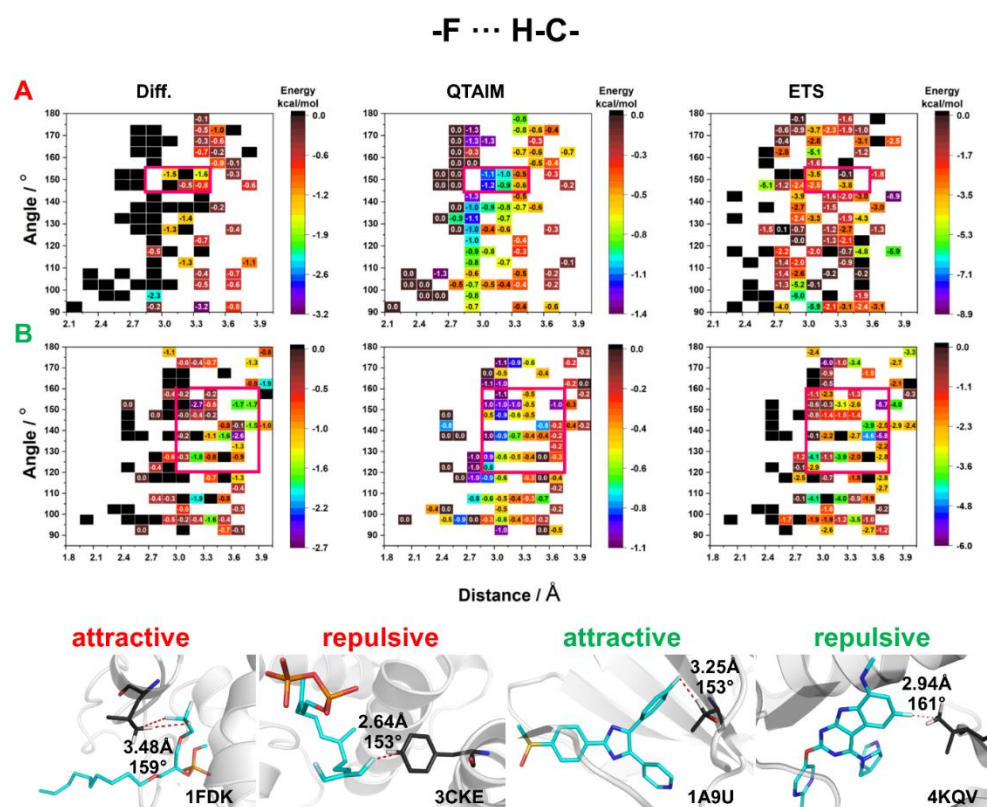


Figure 9. HB energy maps generated based on Diff, QTAIM, and ETS calculations of interaction energy between CH donor and fluorine attached to (A) an aliphatic fragment and (B) an aromatic ring for specific geometric parameters. The areas for which the highest stabilizing energy was observed in all three methods are marked with a red square.

3. Materials and Methods

3.1. PDB Analysis

We performed a statistical analysis of the structural data and investigated in detail the geometric parameters of the intermolecular HBs of fluorine in the structures deposited in the PDB repository. In the first step, all fluorine-containing ligands were identified in the LigandExpo database [21] and then divided into two groups: fluorine attached to an aliphatic carbon (F_{al}) and fluorine attached to an aromatic carbon (F_{ar}). In the next step, all crystal structures containing the abovementioned ligands were identified. The positions of hydrogen atoms were added, considering the stereochemical rules determining the most favorable position of hydrogens, using the Protein Preparation Wizard (Schrödinger Maestro Software) [22]. The in-house script was used to detect the interactions (contact) occurring between fluorine (as an acceptor) in the ligand and the neighboring HB donors (i.e., OH, NH, CH) in the receptor that met the following criteria: a distance of <4 Å and an angle of 90 – 180° .

Based on the obtained data, density maps showing the distribution of the geometric parameters of HBs were generated using R [23] environment as well as RColorBrewer [24], Hexbin [25], Rbokeh [26], and ggplot2 [27] libraries.

3.2. Calculation of Interaction Energy

To determine the energy of the studied intermolecular interactions with fluorine, all complexes obtained from the PDB were divided into subgroups based on the following: (i) angle of HB (ranged from 90° to 180° with a step of 10°), (ii) distance of HB (ranged from 2.5 to 4 Å with a step 0.1 Å), (iii) donor type of HB (OH, NH, ^+NH , CH), and (iv) whether fluorine is bonded to an aromatic (F_{ar}) or aliphatic (F_{al}) carbon. Then, one representative PDB complex was randomly selected from each rectangle defined by unit distance and

angle change. In the next step, all the selected systems were visually inspected to identify those in which the HB with fluorine was not the main stabilizing interaction and the number of adjacent supporting interactions was the smallest. These identified complexes were used to calculate the interaction energy of the fluorine-containing HB for the given geometric parameters. Figure S1 illustrates the distribution of complexes in a given HB distance–angle interval.

The appropriate ionization states at $\text{pH } 7.4 \pm 0.5$ were determined using Epik v3.4 [28,29]. Using an in-house script, the structure of the ligand and the amino acid participating in the interaction was extracted, the missing atoms (included in the peptide bond) were added, and their positions were optimized (OPLS3 force field). The interaction energy of the identified complexes with fluorine was calculated using three commonly used approaches as follows.

The first method, named difference approach or Diff, works based on the assumption that the total interaction energy equals the energy required to separate two interacting molecules. Thus, the energy between the HB donor (X) and the acceptor (Y) is calculated as the difference between the total energy of the $X \cdots Y$ complex and the sum of the total energies of its frozen components [18,30]:

$$E_{int} = E(X \cdot Y) - [E(X) + E(Y)]$$

The energies of the separated molecules, as well as that of the complex, were calculated in Gaussian G16 software [31], using the Minnesota functional M06-2X [32,33] and Karlsruhe basis set def2-TZVP [34]. The polarizable continuum model (PCM) (solvent = water) was used for the calculation [35,36].

The second approach works based on Bader's quantum theory of "atoms in molecules" (QTAIM). In this approach, the topological analysis of electron density was carried out in AIMAll program [37]. The electron density calculated in Gaussian G16 at the M06-2X/def2-TZVP level and the PCM (solvent = water) were used in the analysis. The energy of the noncovalent bonds detected in the crystal structures was calculated using the Espinosa equation as follows:

$$E_{int} = \frac{1}{2}v(r)$$

where E_{int} is the energy of the interatomic interaction and $v(r)$ is the kinetic energy at the bond critical point (BCP). The above equation can be used for all types of HBs, van der Waals interactions, and weak interactions such as $H \cdots H$ and $C-H \cdots O$ [38].

The third approach works based on the energy decomposition analysis. Bonding was analyzed using the extended transition state (ETS) method [39], with the natural orbitals for the chemical valence (NOCV) scheme [40–42]. In this approach, the total energy of bonding between the interacting molecules (ΔE_{int}) is divided into different components as follows:

$$\Delta E_{int} = \Delta E_{dist} + \Delta E_{el} + \Delta E_{Pauli} + \Delta E_{orb} \quad (1)$$

where ΔE_{dist} is the energy required to promote the separated fragments from their equilibrium geometry to the structure they will take up in the complex, ΔE_{el} is the energy of the electrostatic interaction occurring between the two fragments in the supermolecule geometry, ΔE_{Pauli} is the energy of repulsion between the occupied orbitals of the two fragments, and ΔE_{orb} or the orbital interaction term refers to the energy of the stabilizing component due to the final orbital relaxation. All calculations were performed using the Amsterdam Density Functional (ADF) program [43–46], using the ETS-NOCV scheme. The Becke, Lee, Yang, and Parr exchange–correlation functional with the Grimme dispersion correction (B3LYP-D3) was used. A standard double- ζ STO basis containing one set of polarization functions was adopted for all the electrons (TZP). The total (electronic) bonding enthalpies ($\Delta E = \Delta E_{int}$) did not include the zero-point energy (ZPE) additions, finite temperature contributions or basis set superposition error corrections (BSSE).

4. Conclusions

Fluorine is a common substituent in medicinal chemistry and is found in the structure of several currently available blockbuster drugs. This element influences many pharmacokinetic and pharmacodynamic properties of drugs, but its role in stabilizing ligand–biomolecule systems still remains unclear. In this study, we performed a statistical and theoretical analysis of HBs with fluorine found in the PDB database, focusing on the different HB donors (hydroxyl, amine, and methyl groups). The energy range of distinct HBs (i.e., $F \cdots H-O$, $F \cdots H-N^+$, $F \cdots H-N$, $F \cdots H-C$) and optimal ranges of geometric parameters of HBs with fluorine were determined based on the selected PDB complexes.

The results of the analyses showed significant differences in the interaction of fluorine attached to an aliphatic carbon (F_{al}) and fluorine attached to an aromatic carbon (F_{ar}). The $F \cdots H-O$ HBs with F_{ar} are more frequently formed with SER and THR, while those with F_{al} are formed by all amino acids with a polar hydroxyl group. Typically, $F \cdots H-N$ HBs are formed with amino acids that have an amino group in their side chain (ARG, LYS, ASN, GLU). Hydrophobic amino acids most often form $F \cdots H-C$ interactions, which suggests that fluorine prefers a hydrophobic environment in biological systems. It is worth noting that due to the three free electron pairs of fluorine, HBs are only influenced by the donor–acceptor distance and not by the angles. Although the three free electron pairs occupy the entire space around fluorine, $F \cdots H-X$ HBs exhibit the characteristics of HBs, with exceeded standard angles. However, no significant differences were noted in the energies of HBs with fluorine depending on the donor type, which indicates that fluorine acts as a weak HB acceptor for all types of atoms. The optimal ranges of geometric parameters for HBs with fluorine were found to be 150° – 120° and 2.9–3.6 Å. For $F \cdots N^+$ interactions, an HB distance shorter than 2.8 Å showed a destabilizing character in almost 70% of the cases.

It must be emphasized that all the analyzed crystal structures may not be crystallized at the lowest free energy form, and hence the observed interactions might not be in optimal geometries [47]. However, the results suggest that HBs with fluorine are forced to form, due to the stronger ligand–receptor neighboring interactions, which make fluorine the “donor’s last resort” [48]. This is in line with Margareth Etter’s rule that stronger HBs form first, and weaker donors and acceptors interact afterward [47]. All these findings suggest that fluorine does not form strong, stabilizing intermolecular interactions, and thus it seems that indirect influence of this element (electrostatic, inductive, and resonance effects) has a greater impact on the biological activity of compounds than his influences on the pharmacodynamics. The results of this study may contribute to a thorough understanding of hydrogen bonding with fluorine in biological systems which may serve to improve the tools currently available for the rational design of new fluorinated drugs.

Supplementary Materials: The following supporting information can be downloaded, Figure S1: Points of geometrical parameters of hydrogen bonds for which the energy value was calculated; Figure S2: Box plots showing the distribution of the different energy components: kinetic (red), electrostatic (cyan), Coulomb (blue), XC (olive), dispersion (magenta) energy for HB donors and fluorine attached to (A) an aliphatic fragment and (B) an aromatic ring. The energy values were calculated for specific geometric parameters using ETS-NOCV approach. Table S1: Correlation coefficient values calculated between results from every method; Table S2: Pearson test values calculated between results from every method.

Author Contributions: W.P.: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Resources, Writing—original draft, Visualization, Project administration; R.K. (Rafał Kurczab): Conceptualization, Resources, Writing—original draft, Supervision, Project administration; R.K. (Rafał Kafel): Software, Visualization; A.J.B.: Investigation, Writing—original draft, Funding acquisition. All authors have read and agreed to the published version of the manuscript.

Funding: The authors acknowledge the financial support from the National Science Centre, Poland (grant no. 2019/35/N/NZ7/04312) and the statutory funding from the Maj Institute of Pharmacology, Polish Academy of Sciences, Poland.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data were obtained from the PDB repository and the DrugCentral 2021 database (accessed 30 April 2021).

Acknowledgments: Numerical simulations were performed by PLGrid Infrastructure (Prometheus, ACC Cyfronet, AGH). WP acknowledges the support of InterDokMed project no. POWR.03.02.00-00-1013/16. The authors thank Justyna Kalinowska-Thućik (Jagiellonian University), Mariusz Mitoraj (Jagiellonian University), and Filip Sagan (Jagiellonian University) for providing technical and substantive assistance in research.

Conflicts of Interest: The authors declare no conflict of interest.

Sample Availability: Samples of the compounds are not available from the authors.

References

1. Swallow, S. Fluorine in Medicinal Chemistry. In *Progress in Medicinal Chemistry*; Lawton, G., Witty, D.R., Eds.; Elsevier: Amsterdam, The Netherlands, 2015; Volume 54, pp. 65–133.
2. Fried, J.; Sabo, E.F. 9 α -Fluoro derivatives of cortisone and hydrocortisone. *J. Am. Chem. Soc.* **1954**, *76*, 1455–1456. [[CrossRef](#)]
3. Zhou, Y.; Wang, J.; Gu, Z.; Wang, S.; Zhu, W.; Aceña, J.L.; Soloshonok, V.A.; Izawa, K.; Liu, H. Next Generation of Fluorine-Containing Pharmaceuticals, Compounds Currently in Phase II–III Clinical Trials of Major Pharmaceutical Companies: New Structural Trends and Therapeutic Areas. *Chem. Rev.* **2016**, *116*, 422–518. [[CrossRef](#)]
4. Tressaud, A. Fluorine, a key element for the 21st century. In *Fluorine*; Tressaud, A., Ed.; Elsevier: Amsterdam, The Netherlands, 2018; pp. 77–150.
5. Brown, D.G.; Wobst, H.J. A Decade of FDA-Approved Drugs (2010–2019): Trends and Future Directions. *J. Med. Chem.* **2021**, *64*, 2312–2338. [[CrossRef](#)] [[PubMed](#)]
6. Ursu, O.; Holmes, J.; Bologa, C.G.; Yang, J.J.; Mathias, S.L.; Stathias, V.; Nguyen, D.T.; Schürer, S.; Oprea, T. DrugCentral 2018: An update. *Nucleic Acids Res.* **2019**, *47*, D963–D970. [[CrossRef](#)] [[PubMed](#)]
7. Chen, D.; Oezguen, N.; Urvil, P.; Ferguson, C.; Dann, S.M.; Savidge, T.C. Regulation of protein-ligand binding affinity by hydrogen bond pairing. *Sci. Adv.* **2016**, *2*, e1501240. [[CrossRef](#)] [[PubMed](#)]
8. Bitencourt-Ferreira, G.; Veit-Acosta, M.; de Azevedo, W.F. Hydrogen bonds in protein-ligand complexes. *Methods Mol. Biol.* **2019**, *2053*, 93–107. [[PubMed](#)]
9. Sarkhel, S.; Desiraju, G.R. Hydrogen Bonds in Protein-Ligand Complexes: Strong and Weak Interactions in Molecular Recognition. *Proteins Struct. Funct. Genet.* **2004**, *54*, 247–259. [[CrossRef](#)] [[PubMed](#)]
10. Hogendorf, A.S.; Hogendorf, A.; Popiołek-Barczyk, K.; Ciechanowska, A.; Mika, J.; Satała, G.; Walczak, M.; Latacz, G.; Handzlik, J.; Kieć-Kononowicz, K.; et al. Fluorinated indole-imidazole conjugates: Selective orally bioavailable 5-HT7 receptor low-basicity agonists, potential neuropathic painkillers. *Eur. J. Med. Chem.* **2019**, *170*, 261–275. [[CrossRef](#)]
11. Howard, J.A.K.; Hoy, V.J.; O'Hagan, D.; Smith, G.T. How good is fluorine as a hydrogen bond acceptor? *Tetrahedron* **1996**, *52*, 12613–12622. [[CrossRef](#)]
12. Pietruś, W.; Kurczab, R.; Kafel, R.; Machalska, E.; Kalinowska-Thućik, J.; Hogendorf, A.; Żylewski, M.; Baranska, M.; Bojarski, A.J. How can fluorine directly and indirectly affect the hydrogen bonding in molecular systems?—A case study for monofluoroanilines. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2021**, *252*, 119536. [[CrossRef](#)] [[PubMed](#)]
13. Pietruś, W.; Kurczab, R.; Kalinowska-Thućik, J.; Machalska, E.; Golonka, D.; Barańska, M.; Bojarski, A. Influence of fluorine substitution on nonbonding interactions in selected para-halogeno anilines. *ChemPhysChem* **2021**. [[CrossRef](#)]
14. Zhang, G.; He, W.; Chen, D. On difference of properties between organic fluorine hydrogen bond C-H F-C and conventional hydrogen bond. *Mol. Phys.* **2014**, *112*, 1736–1744. [[CrossRef](#)]
15. Meanwell, N.A. Fluorine and Fluorinated Motifs in the Design and Application of Bioisosteres for Drug Design. *J. Med. Chem.* **2018**, *61*, 5822–5880. [[CrossRef](#)]
16. Dunitz, J.D.; Taylor, R. Organic Fluorine Hardly Ever Accepts Hydrogen Bonds. *Chem. A Eur. J.* **1997**, *3*, 89–98. [[CrossRef](#)]
17. Ferreira De Freitas, R.; Schapira, M. A systematic analysis of atomic protein-ligand interactions in the PDB. *Medchemcomm* **2017**, *8*, 1970–1981. [[CrossRef](#)] [[PubMed](#)]
18. Afonin, A.V.; Vashchenko, A.V.; Sigalov, M.V. Estimating the energy of intramolecular hydrogen bonds from ¹H NMR and QAIM calculations. *Org. Biomol. Chem.* **2016**, *14*, 11199–11211. [[CrossRef](#)] [[PubMed](#)]
19. Šolínová, V.; Kašička, V. Determination of acidity constants and ionic mobilities of polyprotic peptide hormones by CZE. *Electrophoresis* **2013**, *34*, 2655–2665. [[CrossRef](#)] [[PubMed](#)]
20. Graton, J.; Besseau, F.; Brossard, A.M.; Charpentier, E.; Deroche, A.; Le Questel, J.Y. Hydrogen-bond acidity of OH groups in various molecular environments (phenols, alcohols, steroid derivatives, and amino acids structures): Experimental measurements and density functional theory calculations. *J. Phys. Chem. A* **2013**, *117*, 13184–13193. [[CrossRef](#)] [[PubMed](#)]
21. Berman, H.; Henrick, K.; Nakamura, H. Announcing the worldwide Protein Data Bank. *Nat. Struct. Mol. Biol.* **2003**, *10*, 980. [[CrossRef](#)] [[PubMed](#)]

22. Madhavi Sastry, G.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W. Protein, and ligand preparation: Parameters, protocols, and influence on virtual screening enrichments. *J. Comput. Aided. Mol. Des.* **2013**, *27*, 221–234. [CrossRef]
23. Team, R.C. Computational Many-Particle Physics. In *Lecture Notes in Physics*; Fehske, H., Schneider, R., Weiße, A., Eds.; Springer: Berlin/Heidelberg, Germany, 2008; Volume 739.
24. Neuwirth, E. *RColorBrewer: ColorBrewer Palettes*; R Package Version 1.1.2; R Package: Vienna, Austria, 2014. Available online: <https://CRAN.R-project.org/package=RColorBrewer> (accessed on 29 December 2021).
25. Lewin-Koh, N. *Hexagon Binning: An Overview*; R Core Team: Vienna, Austria, 2014.
26. Bokeh Development Team. Bokeh: Python Library for Interactive Visualization. 2014. Available online: <http://www.bokeh.pydata.org> (accessed on 29 December 2021).
27. Villanueva, R.A.M.; Chen, Z.J. ggplot2: Elegant Graphics for Data Analysis (2nd ed.). *Meas. Interdiscip. Res. Perspect.* **2019**, *17*, 160–167. [CrossRef]
28. Greenwood, J.R.; Calkins, D.; Sullivan, A.P.; Shelley, J.C. Towards the comprehensive, rapid, and accurate prediction of the favorable tautomeric states of drug-like molecules in aqueous solution. *J. Comput. Aided. Mol. Des.* **2010**, *24*, 591–604. [CrossRef] [PubMed]
29. Shelley, J.C.; Cholleti, A.; Frye, L.L.; Greenwood, J.R.; Timlin, M.R.; Uchimaya, M. Epik: A software program for pKa prediction and protonation state generation for drug-like molecules. *J. Comput. Aided. Mol. Des.* **2007**, *21*, 681–691. [CrossRef]
30. Jabłoński, M. A Critical Overview of Current Theoretical Methods of Estimating the Energy of Intramolecular Interactions. *Molecules* **2020**, *25*, 5512. [CrossRef] [PubMed]
31. Frisch, M.J.; Trucks, G.W.; Schlegel, H.B.; Scuseria, G.E.; Robb, M.A.; Cheeseman, J.R.; Scalmani, G.; Barone, V.; Petersson, G.A.; Nakatsuji, H.; et al. *Gaussian 16 Revision C.01*; Gaussian Inc.: Wallingford, CT, USA, 2016.
32. Zhao, Y.; Truhlar, D.G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: Two new functionals and systematic testing of four M06-class functionals and 12 other functions. *Theor. Chem. Acc.* **2008**, *120*, 215–241. [CrossRef]
33. Lin, Y.-S.; Li, G.-D.; Mao, S.-P.; Chai, J.-D. Long-Range Corrected Hybrid Density Functionals with Improved Dispersion Corrections. *J. Chem. Theory Comput.* **2013**, *9*, 263–272. [CrossRef]
34. Pritchard, B.P.; Altarawy, D.; Didier, B.; Gibson, T.D.; Windus, T.L. New Basis Set Exchange: An Open, Up-to-Date Resource for the Molecular Sciences Community. *J. Chem. Inf. Model.* **2019**, *59*, 4814–4820. [CrossRef]
35. Mennucci, B.; Cammi, R.; Tomasi, J. Analytical free energy second derivatives with respect to nuclear coordinates: Complete formulation for electrostatic continuum solvation models. *J. Chem. Phys.* **1999**, *110*, 6858–6870. [CrossRef]
36. Miertuš, S.; Scrocco, E.; Tomasi, J. Electrostatic interaction of a solute with a continuum. A direct utilization of AB initio molecular potentials for the prevision of solvent effects. *Chem. Phys.* **1981**, *55*, 117–129. [CrossRef]
37. Keith, T.A. *AIMAll*; TK Gristmill Software: Overland Park, KS, USA, 2015.
38. Espinosa, E.; Molins, E.; Lecomte, C. Hydrogen bond strengths revealed by topological analyses of experimentally observed electron densities. *Chem. Phys. Lett.* **1998**, *285*, 170–173. [CrossRef]
39. Ziegler, T.; Rauk, A. On the calculation of bonding energies by the Hartree Fock Slater method—I. The transition state method. *Theor. Chim. Acta* **1977**, *46*, 1–10. [CrossRef]
40. Michalak, A.; Mitoraj, M.; Ziegler, T. Bond orbitals from chemical valence theory. *J. Phys. Chem. A* **2008**, *112*, 1933–1939. [CrossRef] [PubMed]
41. Mitoraj, M.; Michalak, A. Applications of natural orbitals for chemical valence in a description of bonding in conjugated molecules. *J. Mol. Model.* **2008**, *14*, 681–687. [CrossRef] [PubMed]
42. Mitoraj, M.; Michalak, A. Natural orbitals for chemical valence as descriptors of chemical bonding in transition metal complexes. *J. Mol. Model.* **2007**, *13*, 347–355. [CrossRef]
43. Velde, G.; Bickelhaupt, F.M.; Baerends, E.J.; Fonseca Guerra, C.; van Gisbergen, S.J.A.; Snijders, J.G.; Ziegler, T. Chemistry with ADF. *J. Comput. Chem.* **2001**, *22*, 931–967. [CrossRef]
44. Baerends, E.J.; Ros, P. Self-consistent molecular Hartree-Fock-Slater calculations II. The effect of exchange scaling in some small molecules. *Chem. Phys.* **1973**, *2*, 52–59. [CrossRef]
45. Baerends, E.J.; Ellis, D.E.; Ros, P. Self-consistent molecular Hartree-Fock-Slater calculations I. The computational procedure. *Chem. Phys.* **1973**, *2*, 41–51. [CrossRef]
46. Velde, G.; Baerends, E.J. Numerical integration for polyatomic systems. *J. Comput. Phys.* **1992**, *99*, 84–98. [CrossRef]
47. Etter, M.C. Hydrogen bonds as design elements in organic chemistry. *J. Phys. Chem.* **1991**, *95*, 4601–4610. [CrossRef]
48. Taylor, R. The hydrogen bond between N-H or O-H and organic fluorine: Favourable yes, competitive no. *Acta Crystallogr. Sect. B Struct. Sci. Cryst. Eng. Mater.* **2017**, *73*, 474–488. [CrossRef]