Fluorine local environment: from screening to drug design

Anna Vulpetti[1] and Claudio Dalvit[2]

[1] Novartis Institutes for Biomedical Research, Global Discovery Chemistry, CADD, CH 4002 Basel, Switzerland
[2] University of Neuchâtel, Department of Chemistry, CH 2000 Neuchâtel, Switzerland

Fluorine is widely used in the lead optimization phase of drug discovery projects. More recently, fluorine NMR-based spectroscopy has emerged as a versatile, reliable and efficient tool for performing binding and biochemical assays. Different libraries of fluorinated compounds, designed by maximizing the chemical space around the fluorine atom, are screened for identifying binding fragments and for detecting putative fluorophilic hot spots on the desired macromolecular target. A statistical analysis of the fluorine NMR chemical shift, which is a marker of the fluorine local environment, and of the X-ray structures of fluorinated molecules has resulted in the development of the 'rule of shielding'. This method could become a useful tool for lead optimization and for designing novel chemical scaffolds that recognize distinct protein structural motifs.

The fluorine atom is very popular during lead optimization in drug discovery. The outstanding significance of fluorine in the development of bioactive molecules has been reviewed [1–6]. Fluorine is the most abundant halogen in earth crust, but only about a dozen of fluorine-containing natural compounds are known [7]. However, fluorine is frequently used in the pharmaceutical industry and approximately 20–25% of all drugs contain at least one fluorine atom. These include some blockbusters, for example, Prozac (fluoxetine, depression), Celebrex (celecoxib, arthritis), Arcoxia (etoricoxib, arthritis), Sustiva (efavirenz, anti-HIV), Januvia (sitagliptin, diabetes) and Lipitor (atorvastatin, dyslipidemia). The introduction of fluorine in biologically active compounds modulates electronic, lipophilic, steric parameters, which can critically influence pharmacokinetic (PK) and/or pharmacodynamic (PD) properties.

The exchange of proton with fluorine is frequently used to enhance the *metabolic stability* of the molecule owing to the higher dissociation energy of a C–F bond versus a C–H bond (the dissociation energies of C–H and C–F bonds are 98.8 kcal/mol and 105.4 kcal/mol, respectively) [8].

Corresponding authors:. Vulpetti, A. (anna.vulpetti@novartis.com), Dalvit, C. (claudio.dalvit@unine.ch)

Fluorine usually increases *acidity* of acids [8], and decreases *basicity* of bases [9]. For example, an increased acidity of $CF_3SO_2NH_2$ ($pK_a = 5.8$) versus $CH_3SO_2NH_2$ ($pK_a = 10.8$) explains its better affinity towards carboxyanhydrase II ($Ki = 10^{-4}$ M versus $Ki = 2 \times 10^{-3}$M) as the deprotonated NH_2 coordinates the Zn^{2+} [10]. A decrease in the pK_a value is observed upon fluorine introduction in the series $CH_3CH_2NH_2$ ($pK_a = 10.7$), $FCH_2CH_2NH_2$ ($pK_a = 9.0$), $F_2CHCH_2NH_2$ ($pK_a = 7.3$) and $F_3CCH_2NH_2$ ($pK_a = 5.7$). Useful predictive rules have been developed at Roche for tuning the pK_a values of basic amine centers through σ-transmission effects of fluorine, oxygen, nitrogen and sulfur functionalities [11].

Lipophilicity is usually increased if fluorine is in the vicinity of basic nitrogens and if it is introduced to aromatic rings [12]. It is also increased by both trifluoromethyl and trifluoromethoxy substituents, but single fluorine atoms might alter this parameter in either direction. For example, it has been reported that lipophilicity is reduced if fluorine is close in space to oxygen, and if it is on a saturated alkyl group.

Permeability can be increased by the introduction of fluorine in a specific position of a lead molecule [13,14]. In particular, the formation of intramolecular H-bond between a fluorine atom and polar donor functionalities can result in a reduced polar surface area, with a consequent increase in permeability.

1359-6446/06/$ - see front matter http://dx.doi.org/10.1016/j.drudis.2012.03.014

The introduction of fluorine can also be used to enhance *binding affinity* through direct contact with the protein (by means of H-bond interactions, lipophilic contacts and multipolar and/or antiparallel polar interactions), indirect contact with the protein mediated by water molecules [15], or by preorganizing the ligand conformation as needed for the protein recognition [16].

Finally, the fluorine has been used in *NMR spectroscopy* for drug discovery applications. ^{19}F NMR spectroscopy has been shown to be a powerful and sensitive method for studying the metabolism of fluorinated drugs [17] and for performing binding and biochemical assays [18], known as fluorine chemical shift anisotropy and exchange for screening (FAXS) [19,20] and n-fluorine atoms for biochemical screening (n-FABS) [21–25], respectively. ^{19}F NMR-based assays are becoming increasingly popular owing to the versatility, robustness and reliability of the technique and several companies and academic institutions are now investing in this methodology.

In this article we present the combination of computational chemistry, ^{19}F NMR binding assay and X-ray as a powerful and efficient practice in the use of fluorine in drug discovery. Not all fluorine atoms are identical and, depending on the topology around the fluorine (i.e. the different chemical local environment in which the fluorine is embedded), both its ^{19}F NMR chemical shift and its type of interaction with the protein are affected. Thus, a computational procedure was developed to encode the different fluorine local environment in a specific fingerprint descriptor, which is correlated to the fluorine NMR isotropic chemical shifts, which in turn is a good indicator of the type of interactions the fluorine is prone to make with the protein. The analysis of protein–ligand X-ray structures present in the Protein Data Bank (PDB) [26] and the X-ray structures of fluorinated molecules in the Cambridge Structure Database (CSD) [27] support the proposed correlation between the ^{19}F NMR chemical shift and the amphiphatic character of fluorine (i.e. its dual property to act as hydrogen bond acceptor or as hydrophobic moiety).

^{19}F NMR-based binding assay and library design

Fragment-based drug discovery (FBDD) represents an efficient approach for hit identification and optimization. FBDD is now widely recognized for its impact on drug discovery projects and is strongly endorsed by industry in addition to academia. Recently the US FDA has approved the first drug [28] for which there is a clear statement that the drug originated from a fragment screening project initiated in 2005 [29]. A large variety of biophysical techniques can be applied in the initial phase of FBDD [30]. One of these is ^{19}F NMR ligand-based screening experiments known as FAXS. The approach (Fig. 1) consists of first screening library(ies) of fluorinated compounds (typically containing CF and CF_3) to identify binders (direct binding assay) [19,20,31–33].

Different types of fluorinated libraries, as depicted in Fig. 1, can be generated by using different selection criteria. The applied strategy for the design of a fluorinated Novartis fragment library, named *LEF*, which takes into account the local environment of fluorine, was recently described [34]. The procedure is based on a newly developed fluorine fingerprint descriptor (F-FP-5). The F-FP-5 descriptor was derived from the topological-torsion descriptor concept [35]. All the paths consisting of one to five bonds and only starting from the fluorine atom or CF_3 moiety (described as a single dummy atom) are enumerated and encoded in fingerprints. Each atom in the path is described by three parameters: (i) type of element, (ii) number of π electrons and (iii) number of heavy neighbors (not including those in the described path). After generation of the fingerprints, the classical approaches of similarity search (dice similarity metric) are applied for clustering the fingerprints. This approach was used for clustering the collected fluorinated fragments from in-house and commercial sources and for selecting a diverse set of fluorinated fragments (Fig. 1).

Compounds with similar local fluorine-environment tend to have similar chemical shifts. Fluorine chemical shifts are strongly modulated by the number of oxygen, nitrogen and halogen atoms and their positions (α, β and γ, i.e. path of 2, 3 and 4 bonds, respectively) with respect to fluorine. However, the fingerprint was extended up to seven bonds to better take into consideration long-range effects on chemical shift. The new extended fluorine fingerprint, together with distance-weighted k-nearest neighbors (KNN) algorithm, enables the ^{19}F prediction of chemical shift of new CF_3- or CF-containing molecules [36].

The hypothesis at the basis of the fragment library design based on the LEF was that fluorines are not identical and that their chemical environments are relevant for the type of interactions the fluorines make with the protein. Thus, in addition to classical criteria used for designing fragment libraries (maximal global diversity, 2D filters, possible rapid chemical elaboration), all the different chemical topologies around fluorine are represented in the LEF library.

The LEF library has now been expanded by using a diversity oriented synthesis approach [37]. Fragments are synthesized starting from selected scaffolds with or without fluorine and with one to two sites of diversity amenable for parallel chemistry (LEF-DOS) (Fig. 1). The scaffolds are decorated with building blocks with/without containing fluorine depending, respectively, on the absence or presence of fluorine on the scaffold.

The synthesis of *ad hoc* designed fluorinated fragments enables the generation of fluorinated fragments under-represented in the in-house and/or vendor collections. In particular, it gives the opportunity to enhance the fragment chemical space with more 3D-like fragments [38]. The availability of close analogs in the LEF-DOS screening library makes an initial structure–activity relationship (SAR) possible. Moreover, if binders are found, an additional focused small library can be rapidly generated in the follow-up chemistry phase.

Another fluorinated library can be constructed by *defragmenting* known ligands. The introduction of fluorine tags (e.g. CF or CF_3 moieties) in the fragments originated from the defragmentation of known ligand(s) can be carried out by *ad hoc* synthesis or by selecting fluorinated analogs in the in-house and/or commercial collections (Fig. 1).

Finally, a library of *fluorinated peptides* and/or *fluorinated peptide mimetics* of various lengths and amino acid composition can also be generated. The fluorine chemical shift for different fluorinated amino acids and even the fluorine chemical shift of the same fluorinated amino acid present in peptides with different amino acid composition and/or length are distinct enabling the generation of large mixtures. This library is particularly relevant for the screening of inhibitors of protein–protein interactions.

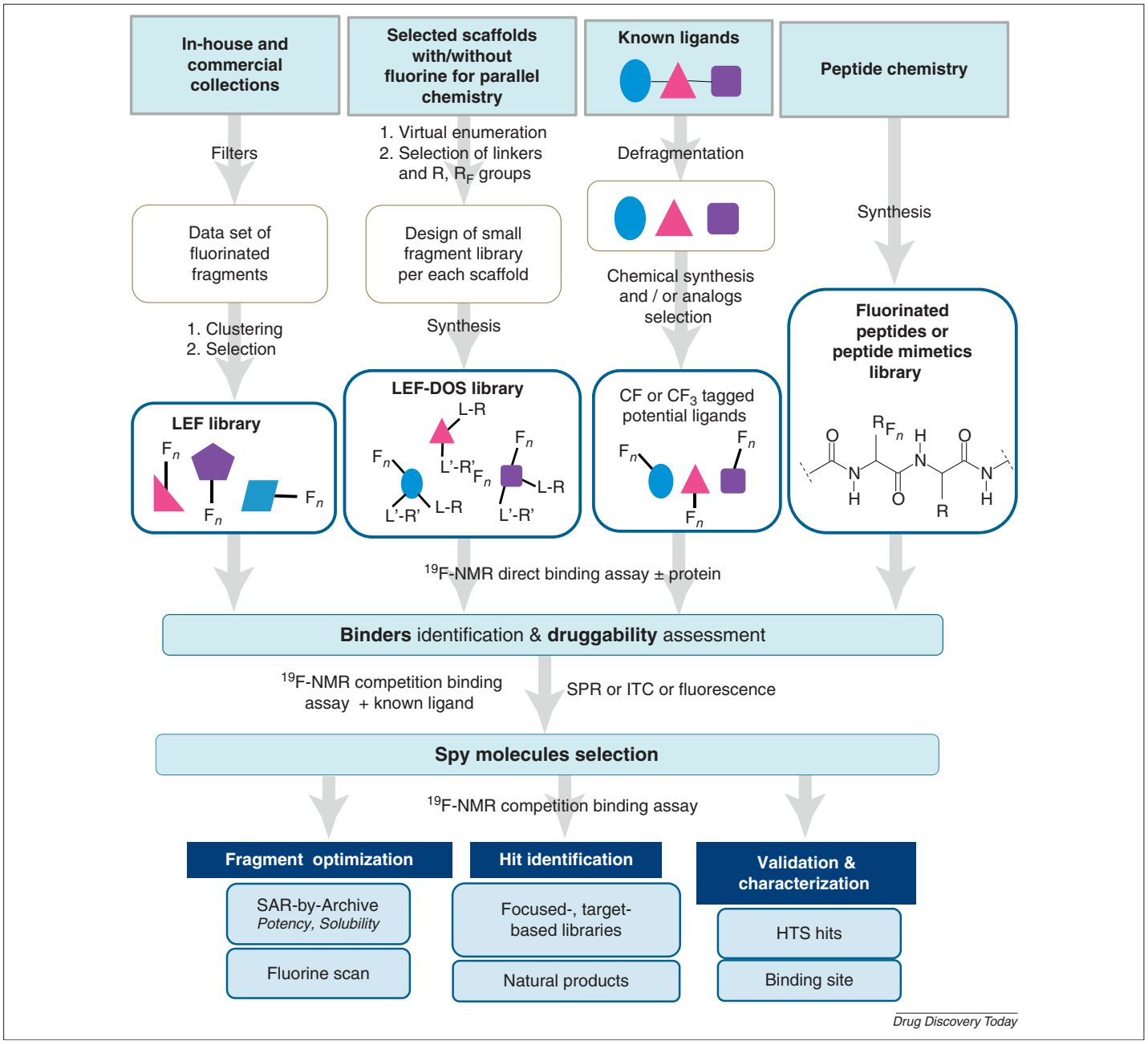


FIGURE 1

Workflow for the screening with the ^{19}F NMR-based binding assay. *Abbreviations*: HTS: high-throughput screening; ITC: isothermal titration calorimetry; LEF: local environment of fluorine; LEF-DOS: local environment of fluorine-diversity oriented synthesis; SAR: structure–activity relationship; SPR: surface plasmon resonance.

The molecules of the fluorine libraries are tested in large mixtures and at low concentrations (10–50 μM). The CF_3 signals have a threefold higher sensitivity compared with the CF signals. Therefore the CF_3-containing molecules can be tested at lower concentrations. However, the CF-containing molecules usually have a higher relative sensitivity to protein binding owing to the large observed difference in chemical shift of the fluorine signal between the free and bound state. Typically, few thousand molecules can be screened per day and a low amount of protein is required for the screening [34] (e.g. 1 mg for a 30 kDa protein is required for screening few thousand compounds).

Molecules of the fluorinated library binding to a protein can be easily identified by inspecting the ^{19}F NMR spectra recorded in the absence and in the presence of the protein. The binding molecules are detected by the reduction in signal intensity or even disappearance of their ^{19}F NMR resonances in the spectrum recorded in the presence of the protein. The deconvolution is performed on-the-fly because each ^{19}F resonance is assigned to a specific molecule in the mixture. Validation of these hits is then performed by the addition of known ligand(s), when available.

Moreover, the screening of the fluorinated libraries is particularly suitable for assessing *ligandability (druggability) of new targets*. Other companies (Abbott [39] and AstraZeneca [40]) have already used fragment based screening (FBS) as an indicator of ligandability. The possibility of screening large mixtures (e.g. >30 compounds per mixture) enables a substantial reduction of protein

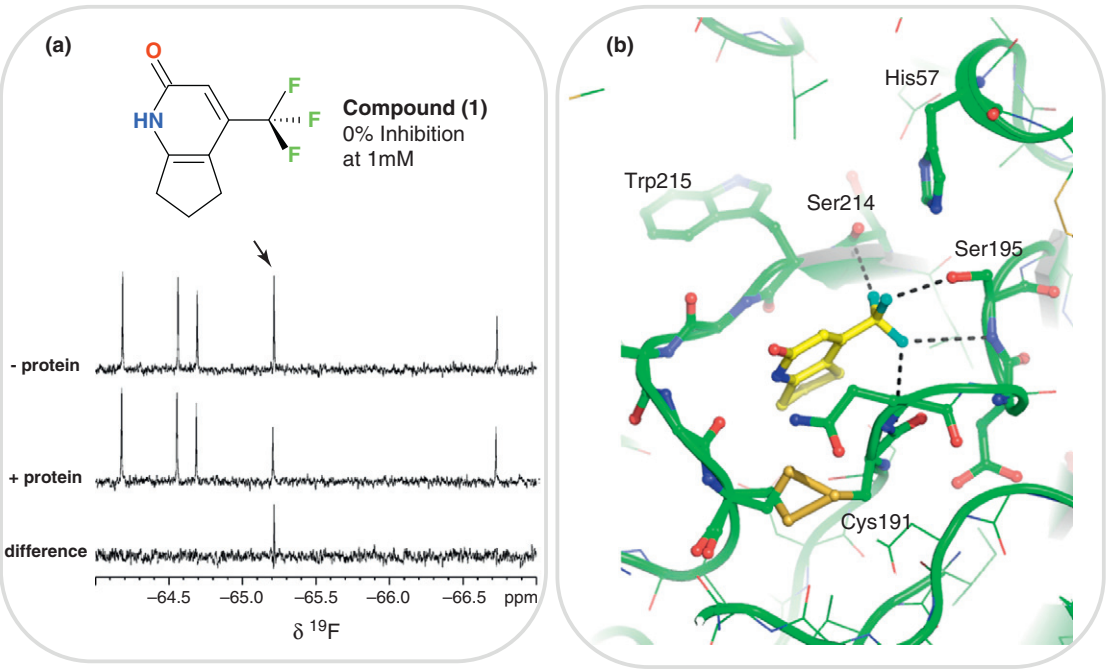


FIGURE 2

(a) ^{19}F NMR-based binding assay performed with the LEF library against bovine trypsin. Only the signal of the molecule interacting with the protein is visible in the difference spectrum (obtained by subtracting the spectra in the absence and presence of the protein) and its chemical shift enables on-the-fly identification of fragment **(1)** as binder. **(b)** High resolution (1.15 Å) X-ray structure (PDB code: 3NK8) of the NMR identified ligand in complex with bovine trypsin [42]. The short interactions of the CF_3 group with the protein are displayed with black dashed lines. *Abbreviations*: LEF: local environment of fluorine; PDB: Protein Database Bank.

consumption and rapid data collection. Therefore even proteins which are difficult to be expressed or handled (e.g. membrane proteins) can be tackled and only a limited protein production effort is necessary for these initial feasibility studies. An example of ^{19}F NMR-based fragment screening as a quick and efficient means of assessing target druggability has been recently reported [41].

The technique has a large dynamic range and thus can detect weak affinity ligands [18,30,34]. This is demonstrated in Fig. 2 which shows the identification of fragment **(1)** as a binder to trypsin [42]. The fragment was inactive in an enzymatic assay performed at one millimolar fragment concentration. Despite the extremely weak binding affinity, the fragment was detected with the FAXS experiment and the X-ray structure of the fragment–protein complex was solved [42].

The validated binders identified in the direct binding assay can then be used as 'spy molecules' for performing competition binding assays for subsequent screening of compounds containing (or not) fluorine atom(s) and for measuring the binding constant of the identified actives.

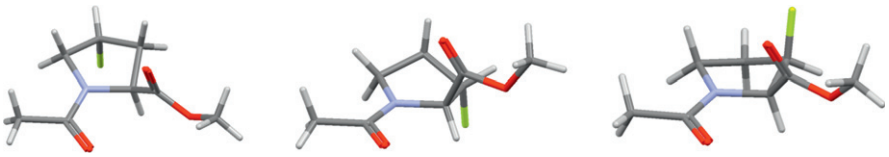
Application of ^{19}F NMR-based competition binding assays include (see Fig. 1):

(i) New *hit finding campaign* by screening additional available libraries of compounds, [e.g. specifically designed for the target under screening (focused-based library) or specifically designed for the family to which the target under screening belongs (target-based library e.g. serine protease library, metal containing protein library, kinase library, protein–protein interaction library, among others)] or natural product extracts. The absence of fluorine moieties in natural products makes the ^{19}F NMR competition binding assay attractive for screening mixtures derived by complex natural extracts or fractions.

(ii) *Hit list triaging* [i.e. high-throughput screening (HTS), hit validation and characterization]. The use of ^{19}F NMR experiment has been reported in the context of a HTS triage effort [43,44]. In one case two CF_3-containing fragments were used to probe inhibitor binding sites on an essential enzyme target within the coenzyme A biosynthetic pathway, phosphopantetheine adenylyltransferase (PPAT) [44]. The two weakly binding fragments were used as spy probes to classify HTS hits as adenosine-5′-triphosphate (ATP) site binders versus no ATP site binders.

(iii) *SAR-by-Archive* (hit optimization) around the identified fluorinated binders. At this stage close analogs or larger molecules [45] that contain the identified binding fragment(s) are collected from in-house archive or purchased. The method enables qualitative (ranking) or quantitative (K_I determination) SAR. It should be pointed out that at this stage the selected molecules for screening do not require the presence of fluorine. The aim is to improve potency, in addition to solubility, to enhance the success of protein–ligand crystal structure determination.

(iv) *Fluorine scan* (i.e. the testing of isomers or close analogs differing in the position of the fluorine moiety) to enable the identification of the possible presence of a fluorophilic protein environment.



(4R)-F-proline
C_4-exo (F_{axial})
RISDEC

(3S)-F-proline
C_3-exo (F_{axial})
LEWXER

(3R)-F-proline
C_3-endo (F_{axial})
LEWPAF

SCHEME 1

Cambridge Structure Database crystal structures of 3- and 4-fluoroproline derivatives: RISDEC [64], LEWXER and LEWPAF [65].

Fluorophilic protein environments

The use of a fluorine scan on tricyclic inhibitors of thrombin and on benzimidazole inhibitors of neprilysin [46,47] enabled the map of the fluorophilicity and/or fluorophobicity of the enzyme active sites. Several crystal structures solved for various fluorinated inhibitors in complex with thrombin and neprilysin, in addition to extensive searches in the PDB, revealed the evidence of specific favorable protein–fluorine interactions. Orthogonal multipolar C–F···C=O interactions and C–F···H–C_α close contacts were nicely revealed during the fluorine scan of thrombin inhibitors [46]. Orthogonal multipolar interactions between the C–F and the guanidinium group of arginine were observed during the fluorine scan on neprilysin [47].

The optimized local diversity around fluorine in the *LEF* library combined with the high sensitivity of the ^{19}F NMR ligand-based screening experiments is an efficient approach not only for the initial binding fragments detection, but also for probing the presence of fluorophilic protein environments. An example is the CF_3 motif of compound **(1)** bound to trypsin. The X-ray structure shown in Fig. 2 revealed the involvement of the CF_3 in polar contacts with the catalytic Ser195 and in multipolar interactions with the backbone carbonyl of Ser214 and Cys191 (Fig. 2) [42].

A similar 3D protein environment around a specific fluorine of interest (i.e. making the same number and type of interactions with the protein) can be identified in the PDB using a recently described method [42]. All the protein sub-sites are extracted around each ligand fluorine atom contained in the PDB and in-house complex structures. The protein atoms at a distance smaller than 4.2 Å from each fluorine are assigned to one of eight pharmacophoric features computed considering three main atomic properties (π, donor and acceptor character). Three atom types are included in the eight pharmacophoric features to describe specific interactions involving fluorine. The use of three-point pharmacophoric description (common practice in ligand-based pharmacophoric generation) enables the clustering and the comparison of the sub-pockets in terms of the number and type of common 3D arrangement of features.

For example, the protein environment around the fluorine atoms of compound **(1)** of Fig. 2 [42] used as a query, retrieves similar sub-pockets from structurally related proteins, such as elastase [48], but also structurally unrelated proteins, such as BACE1, LTP1 [49] and 11β-HSD1 [50]. It is interesting to observe that in two cases the bound ligands are fluorinated organic solvents (CF_3CH_2OH, CF_3COOH) suggesting that the sub-pocket under analysis significantly contributes to the binding energy of compound **(1)** (i.e. is a fluorophilic hot spot) and that the use of fluorinated solvents is useful to identify fluorophilic protein environments.

The method is not limited to queries containing a fluorine atom, but it can also be extended to ligands that do not contain fluorine. In this case the virtual query is constructed by replacing *in silico* the proton or other atom types with a fluorine atom at different positions of the ligand in the protein bound crystal structure (virtual fluorine scan) [42]. The identification of similar fluorophilic sub-pockets in other proteins would represent an indication that the chemical substitution of that proton or other

BOX 1

Effect of fluorine on conformational preference.

It is important to point out that the introduction of an electronegative atom (such as O, N or F) in β position to a fluorine atom results in a shielded fluorine (either aromatic or aliphatic). In addition, the conformational preference of alkyl fluorine is also affected.

The classical example is that of 1,2-difluoroalkane which prefers to adopt a gauche conformation (e.g. AMECIE CSD code) [57]. This effect is called the *gauche effect* and it can be explained based on the σ hyperconjugation involving charge transfer from the σ C–H electron donor molecular orbital (HOMO) to the σ* C–F acceptor antibonding molecular orbital (LUMO). This type of orbital interaction is maximized when the σ C–H bonds and the σ* C–F antibonds are in anti-orientation to each other. A density functional theory (DFT) study on the origin of the fluorine gauche effect in different substituted fluoroethanes is reported [58].

Another well-studied case is the effect of fluorine at position 4 of proline in affecting proline pucker conformation at position 4 (C_4). Theoretical studies suggest that the driving force in the preferred puckering is the hyperconjugation interaction (i.e. the donation from the $\sigma(C_3–H)_{ax}$ and $\sigma(C_5–H)_{ax}$ bonding orbitals into the $\sigma^*(C_4–F)_{ax}$ antibonding orbitals). As a consequence, (4*R*)- and (4*S*)-fluoroproline epimers have an opposite proline pucker conformation preference, C_4-*exo* and C_4-*endo*, respectively. Same effect on the proline puckering is also observed by placing the electronegative F atom at position 3 on the proline. Scheme 1 shows three of the available CSD crystal structures on 3- and 4-fluoroproline derivatives. The fluorine substituents that favor the C_4-*exo* pucker, namely (4*R*)- and (3*R*)-fluorine proline result in a preference for the *trans*-peptidic bond of proline [59].

The large C–F dipole drives the C–F bond to be anti-parallel to the C=O bond in α-fluoroamides (FCH_2CONH_2) with a calculated $\Delta E = E_{cis} - E_{trans}$ of +6 kcal/mol. Both the gauche effect between the F and N in β (i.e. F–C–C–N) combined with the α-fluoroamide (dipole–dipole effect) are relevant in controlling peptide conformations (e.g. in β-amino acids *R*-CONH–CH_2–CH(F)–CONH-*R*′). The fluorine stereoelectronic effect has recently emerged as a powerful means to restrict the local chain conformation of polypeptide sequences and to drive asymmetric organocatalysis [60,61].

atom types with fluorine could improve the binding affinity of the starting molecule.

Nature of fluorine–protein interactions

There are various aspects that one needs to consider upon addition of fluorine to a molecule: (i) effect on ligand conformation, (ii) redistribution of electron density owing to polarization of covalent bonds upon fluorination [8,16], (iii) inductive effect on vicinal functional group acidity [8,9], (iv) change in desolvation [51], among others.

Some of the widely described effects on ligand conformation due to fluorine introduction are explained considering the electrostatic and stereoelectronic properties of the C–F bond, which has a large dipole (Box 1).

In addition to the different aspects mentioned above, the ^{19}F NMR isotropic chemical shift parameter has also been proposed for the selection of the appropriate fluorine to be inserted into a molecule [52]. The ^{19}F chemical shift reflects the local environment around the fluorine atom and it is a marker of the character of the C–F bond. Therefore, the dual property of fluorine as a hydrogen bond acceptor and as a hydrophobic moiety can be inferred from the ^{19}F NMR chemical shift.

A correlation between the fluorine chemical shift measured in ^{19}F NMR spectroscopy and the type of fluorine–protein interactions observed in the crystal structures has been found [52]. Shielded fluorines, according to NMR chemical shift (i.e. with increased electron density) are observed preferentially in close contact to hydrogen bond donors of the protein suggesting the

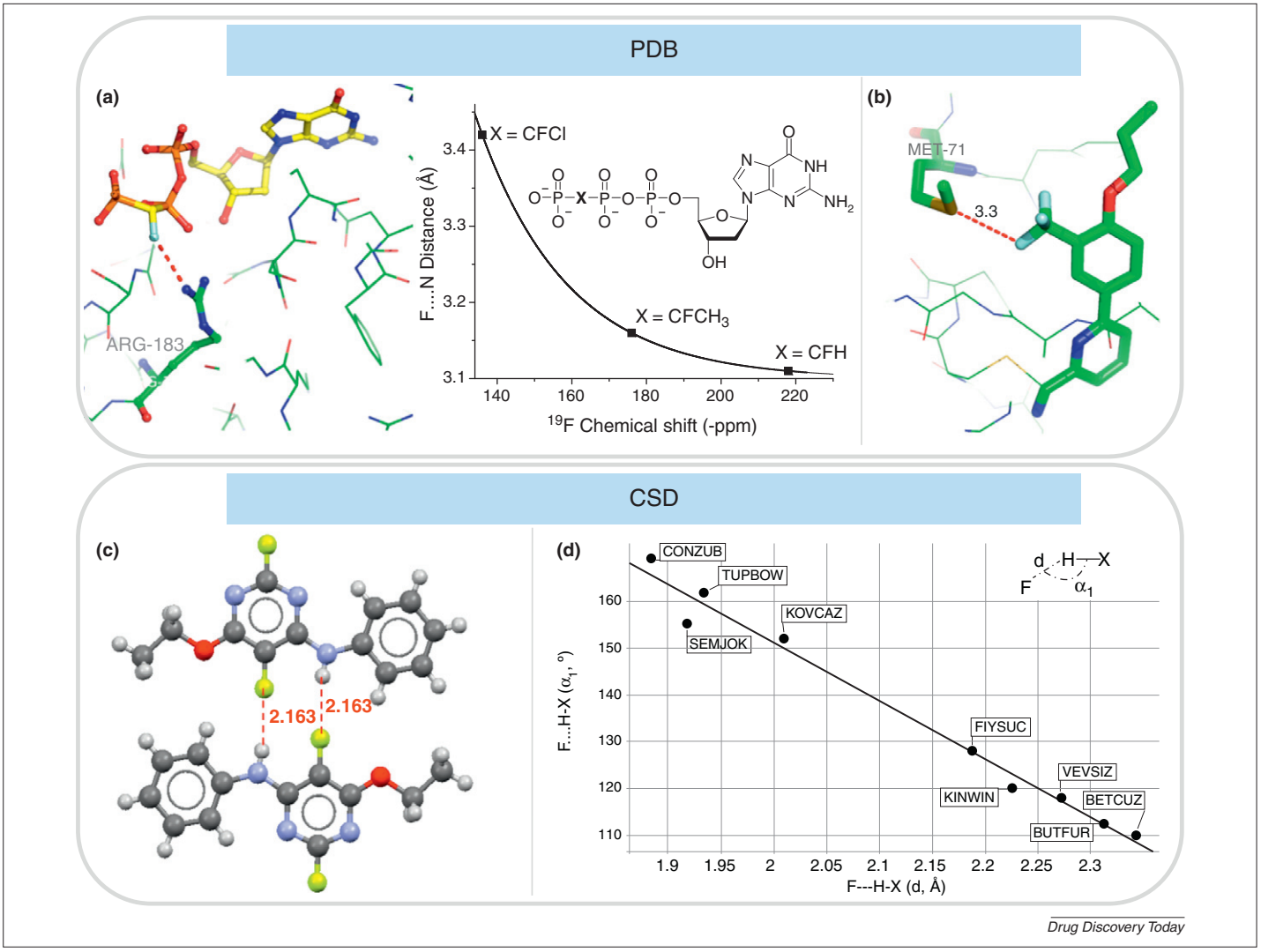


FIGURE 3

(a) X-ray structure of DNA polymerase β in complex with β,γ-CHF-dGTP (PDB code: 2PXI) [53] showing the intermolecular hydrogen bond involving the secondary alkyl fluorine and the NH_2 of Arg183 and the plot of the F···N (Arg183 NH_2) distance for the three displayed ligands in complex with DNA polymerase β [53,54] as a function of their ^{19}F isotropic chemical shifts [52]. **(b)** X-ray structure of a ligand containing a deshielded CF_3 moiety in complex with cathepsin S (PDB code: 3N3G) [62] showing the short intermolecular distance of the sulfur atom of Met71 to the CF_3 moiety. **(c)** X-ray structure of a molecule containing a shielded fluorine atom on a sp^2-hybridized carbon showing the intermolecular hydrogen bond between fluorine and the NH of another molecule present in the same unit cell (CSD code: HUSDEF) [63]. **(d)** Plot of the angle F···H–X (α_1) as a function of the short intermolecular F···H ($\leq$2.35 Å) distance (*d*) for the primary and shielded secondary fluorine atoms extracted from the CSD. The best fit of the data with the straight line $\alpha_1 = 402 - 125*d$ and the CSD codes of the different structures are displayed [56]. *Abbreviations*: CSD: Cambridge Structure Database; PDB: Protein Database Bank.

possibility of intermolecular hydrogen bond formation. An example of a shielded secondary alkyl fluorine involved in hydrogen bond is reported in Fig. 3a. A relationship between the ^{19}F NMR chemical shift and hydrogen bond length, extracted from experimental data [53,54], as shown in the plot of Fig. 3a, and computed with quantum chemical methods [55], has also been derived [52]. Shorter hydrogen bond distances (i.e. stronger hydrogen bonds) are observed with shielded fluorine atoms. On the contrary, deshielded fluorines are found preferentially in close contact with hydrophobic side chains, with the carbon of the carbonyl of the protein backbone and with the sulfur atom of methionines. A CF_3 moiety on aromatic ring with an alkyl *ortho* substitution is significantly deshielded and, as in the example reported in Fig. 3b, is found in close proximity to the sulfur of methionine.

Further evidence for the correlation between ^{19}F NMR isotropic chemical shift and close intermolecular F···H–X (with X = N or O) contacts has been found by analyzing the CSD X-ray structures of fluorinated molecules [56]. Figure 3c shows an example of a shielded fluorine bound to a sp^2-hybridized carbon involved in an intermolecular hydrogen bond with the NH of the other molecule present in the same unit cell. The arrangement of the intermolecular hydrogen bonds in the structure of Fig. 3c is reminiscent of the ten-member rings of anti-parallel β-sheets. In addition, a relationship between F···H distance and F···H–X angle for the primary and shielded secondary alkyl fluorine was also derived (Fig. 3d) with a constant distance F···X of 2.889 ± 0.037 Å. On the basis of all these findings, a 'rule of shielding' [52,56] was proposed which provides some insight and guidelines in the selection of the appropriate fluorinated moiety to be judiciously inserted into the molecule for making the most favorable interactions with the receptor. Several fluorinated scaffolds have been designed based on the rule of shielding for recognizing protein distinct structural motifs.

Concluding remarks

It is fair to conclude that the use of fluorine serves multiple purposes in drug discovery projects from the identification of the initial binding fragments with ^{19}F NMR spectroscopy to their optimization to lead molecules with the judicious introduction of the appropriate fluorinated moiety aimed at improving the affinity, selectivity and physical–chemical properties. The ^{19}F NMR chemical shift contains the information of the environment around the fluorine atom that is of paramount relevance for the interactions of the fluorine moiety with the receptor. It is envisaged that the rule of shielding, derived from these studies, could become a useful tool for lead optimization and for the design of novel chemical scaffolds that recognize protein distinct structural motifs.

References

1 Müller, K. *et al.* (2007) Fluorine in pharmaceuticals: looking beyond intuition. *Science* 317, 1881–1886
2 Bohm, H.J. *et al.* (2004) Fluorine in medicinal chemistry. *ChemBioChem* 5, 637–643
3 Kirk, K.L. (2006) Fluorine in medicinal chemistry: recent therapeutic applications of fluorinated small molecules. *J. Fluorine Chem.* 127, 1013–1029
4 Dolbier, W.R. (2005) Fluorine chemistry at the millennium. *J. Fluorine Chem.* 126, 157–163
5 Hagmann, W.K. (2008) The many roles for fluorine in medicinal chemistry. *J. Med. Chem.* 51, 4359–4369
6 Purser, S. *et al.* (2008) Fluorine in medicinal chemistry. *Chem. Soc. Rev.* 37, 320–330
7 Murphy, C.D. *et al.* (2003) Fluorinated natural products: the biosynthesis of fluoroacetate and 4-fluorothreonine in *Streptomyces cattleya*. *Chemosphere* 52, 455–461
8 O'Hagan, D. (2008) Understanding organofluorine chemistry An introduction to the C–F bond. *Chem. Soc. Rev.* 37, 308–319
9 Schlosser, M. (1998) Parametrization of substituents: effects of fluorine and other heteroatoms on OH, NH, and CH acidities. *Angew. Chem. Int. Ed.* 37, 1496–1513
10 Maren, T.H. and Conroy, C.W. (1993) A new class of carbonic-anhydrase inhibitor. *J. Biol. Chem.* 268, 26233–26239
11 Morgenthaler, M. *et al.* (2007) Predicting properties and tuning physicochemical in lead optimization: amine basicities. *ChemMedChem* 2, 1100–1115
12 Leroux, F. *et al.* (2005) Alpha-fluorinated ethers, thioethers, and amines: anomerically biased species. *Chem. Rev.* 105, 827–856
13 Gerebtzoff, G. *et al.* (2004) Halogenation of drugs enhances membrane binding and permeation. *ChemBioChem* 5, 676–684
14 Kuhn, B. *et al.* (2010) Intramolecular hydrogen bonding in medicinal chemistry. *J. Med. Chem.* 53, 2601–2611
15 Chamberlain, B.T. *et al.* (2012) Stereospecific formation of a ternary complex of (S)-α,β-fluoromethylene-dATP with DNA Pol β. *ChemBioChem* 13, 528–530
16 Hunter, L. (2010) The C–F bond as a conformational tool in organic and biological chemistry. *Beilstein J. Org. Chem.* 6, 1–14
17 Malet-Martino, M. *et al.* (2006) Fluorine nuclear magnetic resonance spectroscopy of human biofluids in the field of metabolic studies of anticancer and antifungal fluoropyrimidine drugs. *Clin. Chim. Acta* 366, 61–73
18 Dalvit, C. (2007) Ligand- and substrate-based F-19 NMR screening: principles and applications to drug discovery. *Prog. Nucl. Magn. Reson. Spectrosc.* 51, 243–271
19 Dalvit, C. *et al.* (2003) Fluorine-NMR experiments for high-throughput screening: theoretical aspects, practical considerations, and range of applicability. *J. Am. Chem. Soc.* 125, 7696–7703
20 Dalvit, C. *et al.* (2002) Fluorine-NMR competition binding experiments for high-throughput screening of large compound mixtures. *Comb. Chem. High Throughput Screen.* 5, 605–611
21 Dalvit, C. *et al.* (2004) Reliable high-throughput functional screening with 3-FABS. *Drug Discov. Today* 9, 595–602
22 Kichik, N. *et al.* (2010) Simultaneous F-19 NMR screening of prolyl oligopeptidase and dipeptidyl peptidase IV inhibitors. *ChemBioChem* 11, 1115–1119
23 Giralt, E. (2011) *In vitro* screening: screening by nuclear magnetic resonance. In *Protein Surface Recognition: Approaches for Drug Discovery* (Giralt, E. *et al.* eds), pp. 237–250, John Wiley & Sons Ltd.
24 Tarrago, T. *et al.* (2006) Identification by F-19 NMR of traditional chinese medicinal plants possessing prolyl oligopeptidase inhibitory activity. *ChemBioChem* 7, 827–833
25 Forino, M. *et al.* (2005) Efficient synthetic inhibitors of anthrax lethal factor. *Proc. Natl. Acad. Sci. U. S. A.* 102, 9499–9504
26 Berman, H.M. *et al.* (2000) The Protein Data Bank. *Nucleic Acids Res.* 28, 235–242
27 Allen, F.H. (2002) The Cambridge Structural Database: a quarter of a million crystal structures and rising. *Acta Crystallogr. B* 58, 380–388
28 Bollag, G. *et al.* (2010) Clinical efficacy of a RAF inhibitor needs broad target blockade in BRAF-mutant melanoma. *Nature* 467, 596–599
29 Tsai, J. *et al.* (2008) Discovery of a selective inhibitor of oncogenic B-Raf kinase with potent antimelanoma activity. *Proc. Natl. Acad. Sci. U. S. A.* 105, 3041–3046
30 Dalvit, C. (2009) NMR methods in fragment screening: theory and a comparison with other biophysical techniques. *Drug Discov. Today* 14, 1051–1057
31 Poppe, L. *et al.* (2007) Discovery of ligands for Nurr1 by combined use of NMR screening with different isotopic and spin-labeling strategies. *J. Biomol. Screen.* 12, 301–311
32 Tengel, T. *et al.* (2004) Use of ^{19}F NMR spectroscopy to screen chemical libraries for ligands that bind to proteins. *Org. Biomol. Chem.* 2, 725–731
33 Klages, J. *et al.* (2006) NMR-based screening: a powerful tool in fragment-based drug discovery. *Mol. Biosyst.* 2, 319–331
34 Vulpetti, A. *et al.* (2009) Design and NMR-based screening of LEF, a library of chemical fragments with different local environment of fluorine. *J. Am. Chem. Soc.* 131, 12949–12959
35 Nilakantan, R. *et al.* (1987) Topological torsion – a new molecular descriptor for SAR applications. Comparison with other descriptors. *J. Chem. Inf. Comput. Sci.* 27, 82–85
36 Vulpetti, A. *et al.* (2010) ^{19}F NMR chemical shift prediction with fluorine fingerprint descriptor. *J. Fluorine Chem.* 131, 570–577

37 Hung, A.W. *et al.* (2011) Route to three-dimensional fragments using diversity-oriented synthesis. *Proc. Natl. Acad. Sci. U. S. A.* 108, 6799–6804
38 Lovering, F. *et al.* (2009) Escape from flatland: increasing saturation as an approach to improving clinical success. *J. Med. Chem.* 52, 6752–6756
39 Hajduk, P.J. *et al.* (2005) Druggability indices for protein targets derived from NMR-based screening data. *J. Med. Chem.* 48, 2518–2525
40 Edfeldt, F.N.B. *et al.* (2011) Fragment screening to predict druggability (ligandability) and lead discovery success. *Drug Discov. Today* 16, 284–287
41 Jordan, J.B. *et al.* (2012) Fragment based drug discovery: practical implementation based on ^{19}F NMR spectroscopy. *J. Med. Chem.* 55, 678–687
42 Vulpetti, A. *et al.* (2010) Combined use of computational chemistry, NMR screening, and X-ray crystallography for identification and characterization of fluorophilic protein environments. *Proteins* 78, 3281–3291
43 Boettcher, A. *et al.* (2010) Fragment-based screening by biochemical assays: systematic feasibility studies with trypsin and MMP12. *J. Biomol. Screen.* 15, 1029–1041
44 Miller, J.R. *et al.* (2010) The use of biochemical and biophysical tools for triage of high-throughput screening hits – a case study with *Escherichia coli* phosphopantetheine adenylyltransferase. *Chem. Biol. Drug Des.* 75, 444–454
45 Crisman, T.J. *et al.* (2008) 'Virtual fragment linking': an approach to identify potent binders from low affinity fragment hits. *J. Med. Chem.* 51, 2481–2491
46 Olsen, J.A. *et al.* (2003) A fluorine scan of thrombin inhibitors to map the fluorophilicity/fluorophobicity of an enzyme active site: evidence for C–F···C═O interactions. *Angew. Chem. Int. Ed.* 42, 2507–2511
47 Morgenthaler, M. *et al.* (2008) A fluorine scan of non-peptidic inhibitors of neprilysin: fluorophobic and fluorophilic regions in an enzyme active site. *J. Fluorine Chem.* 129, 852–865
48 Mattos, C. *et al.* (2006) Multiple solvent crystal structures: probing binding sites, plasticity and hydration. *J. Mol. Biol.* 357, 1471–1482
49 Bakan, B. *et al.* (2009) The crystal structure of oxylipin-conjugated barley LTP1 highlights the unique plasticity of the hydrophobic cavity of these plant lipid-binding proteins. *Biochem. Biophys. Res. Commun.* 390, 780–785
50 Julian, L.D. *et al.* (2008) Discovery of novel, potent benzamide inhibitors of 11β-hydroxysteroid dehydrogenase type 1 (11β-HSD1) exhibiting oral activity in an enzyme inhibition *ex vivo* model. *J. Med. Chem.* 51, 3953–3960
51 Biffinger, J.C. *et al.* (2004) The polar hydrophobicity of fluorinated compounds. *ChemBioChem* 5, 622–627
52 Dalvit, C. and Vulpetti, A. (2011) Fluorine–protein interactions and ^{19}F NMR isotropic chemical shifts: an empirical correlation with implications for drug design. *ChemMedChem* 6, 104–114
53 McKenna, C.E. *et al.* (2007) (R)-beta,gamma-fluoromethylene-dGTP–DNA ternary complex with DNA polymerase beta. *J. Am. Chem. Soc.* 129, 15412–15413
54 Batra, V.K. *et al.* (2010) Halogenated beta,gamma-methylene- and ethylidene-dGTP–DNA ternary complexes with DNA polymerase beta: structural evidence for stereospecific binding of the fluoromethylene analogues. *J. Am. Chem. Soc.* 132, 7617–7625
55 Samsonov, S.A. *et al.* (2009) Fluorine in protein environments: a QM and MD study. *J. Phys. Chem. B* 113, 16400–16408
56 Dalvit, C. and Vulpetti, A. (2012) Intermolecular and intramolecular hydrogen bonds involving fluorine atoms: implications for recognition, selectivity and chemical properties. *ChemMedChem* 7, 262–272
57 Goodman, L. and Sauers, R.R. (2005) 1-Fluoropropane. Torsional potential surface. *J. Chem. Theory Comput.* 1, 1185–1192
58 Buissonneaud, D.Y. *et al.* (2010) A DFT study on the origin of the fluorine gauche effect in substituted fluoroethanes. *Tetrahedron* 66, 2196–2202
59 Salwiczek, M. *et al.* (2012) Fluorinated amino acids: compatibility with native protein structures and effects on protein–protein interactions. *Chem. Soc. Rev.* 41, 2135–2171
60 Sparr, C. *et al.* (2009) The fluorine-iminium ion gauche effect: proof of principle and application to asymmetric organocatalysis. *Angew. Chem. Int. Ed.* 48, 3065–3068
61 Zimmer, L.E. *et al.* (2011) Fluorine conformational effects in organocatalysis: an emerging strategy for molecular design. *Angew. Chem. Int. Ed.* 50, 2–14
62 Cai, J.Q. *et al.* (2010) 4-(3-Trifluoromethylphenyl)-pyrimidine-2-carbonitrile as cathepsin S inhibitors: N3, not N1 is critically important. *Bioorg. Med. Chem. Lett.* 20, 4507–4510
63 Parks, E.L. *et al.* (2010) Trisubstituted pyrimidine derivatives from tetrafluoropyrimidine. *Tetrahedron* 66, 6195–6204
64 Panasik, N. *et al.* (1994) Inductive effects on the structure of proline residues. *Int. J. Pept. Protein Res.* 44, 262–269
65 Kim, W. *et al.* (2006) Fluoroproline flip-flop: regiochemical reversal of a stereoelectronic effect on peptide and protein structures. *Angew. Chem. Int. Ed.* 45, 8141–8145