Fluorine Interactions at the Thrombin Active Site: Protein Backbone Fragments H–C_α–C=O Comprise a Favorable C–F Environment and Interactions of C–F with Electrophiles

Jacob A. Olsen,[b] David W. Banner,*[a] Paul Seiler,[b] Björn Wagner,[a] Thomas Tschopp,[a] Ulrike Obst-Sander,[a] Manfred Kansy,[a] Klaus Müller,*[a] and François Diederich*[b]

In a systematic fluorine scan of a rigid inhibitor to map the fluorophilicity/fluorophobicity of the active site in thrombin, one or more F substituents were introduced into the benzyl ring reaching into the D pocket. The 4-fluorobenzyl inhibitor showed a five to tenfold higher affinity than ligands with other fluorination patterns. X-ray crystal-structure analysis of the protein–ligand complex revealed favorable C–F···H–C_α–C=O and C–F···C=O interactions of the 4-F substituent of the inhibitor with the backbone H–C_α–C=O unit of Asn98. The importance of these interactions was further corroborated by the analysis of small-molecule X-ray crystal-structure searches in the Protein Data Base (PDB) and the Cambridge Structural Database (CSD). In the C–F···C=O interactions that are observed for both aromatic and aliphatic C–F units and a variety of carbonyl and carboxyl derivatives, the F atom approaches the C=O C atom preferentially along the pseudotrigonal axis of the carbonyl system. Similar orientational preferences are also seen in the dipolar interactions C–F···C≡N, C–F···C–F, and C–F···NO_2, in which the F atoms interact at sub-van der Waals distances with the electrophilic centers.

Introduction

Organic fluorine is rarely found in nature, as evidenced by the small number of known fluorinated natural products. In contrast, it is abundant in synthetic materials as documented by the nearly 1.8 million C- and F-containing molecules listed by Chemical Abstract Services (CAS). Moreover, the number of fluorinated drugs is continuously increasing, with 1020 (~5%) included in the Derwent Drug File[1] which contains 21146 chemical entities, and more than 150 fluorinated drugs currently in clinical use.

The introduction of F atoms causes minimal steric effects due to their small size, with a van der Waals radius of 1.47 Å (1.20 Å for H),[2,3] while several important pharmacokinetic properties, such as metabolic stability and absorption properties, can be modulated in a favorable way.

Much less is known about how the thermodynamics of protein–ligand interactions are affected by the replacement of H by F, which is still generally taken as a bioisosteric replacement. Characteristic physical properties of F substituents are a very low polarizability (hard atom), the highest electronegativity in the periodic table, and the presence of three tightly bound nonbonding electron pairs.[2] They confer unique behavior on fluoro-organic compounds: teflon and other fluorinated compounds are neither hydrophilic nor hydrophobic.[4] For a more effective use of organic fluorine in medicinal chemistry, tuning not only pharmacokinetic properties but also ligand binding affinity, a better understanding of the nature of favorable (fluorophilic) and unfavorable (fluorophobic) environments for F substituents is required.

A survey of the literature provides some insight into the nature of favorable F environments in proteins. The C–F unit has been reported to interact favorably with strong H-bond donors, such as the N–H groups of backbone amide linkages and His side chains or the H–O groups of Tyr, Ser, and protein-bound water.[5–8] These contacts are possibly best characterized as C–F···H–X dipolar interactions rather than true H-bonds.[8] Lipophilic side chains constitute another favorable F environment in proteins. Notably, close contacts are observed between aromatic C–H, for example in Phe, and F–C residues.[9,10] In particular, the CF_3 group has a pronounced lipophilicity, as reflected by its Hansch–Leo substituent parameter π of 0.88 ($-CH_3$: $\pi = 0.56$, $-CH_2CH_3$: $\pi = 1.02$).[11–13] In other cases, the effect of F substitution on protein–ligand binding is a more indirect one. Thus, the introduction of F atoms into aromatic

[a] Dr. D. W. Banner, B. Wagner, Dr. T. Tschopp, Dr. U. Obst-Sander, Dr. M. Kansy, K. Müller
Pharma Research Basel, Discovery Chemistry, F. Hoffmann–La Roche Ltd
4070 Basel (Switzerland)
Fax: (+41) 61-688-1745
E-mail: david.banner@roche.com
E-mail: klaus.mueller@roche.com

[b] Dr. J. A. Olsen, P. Seiler, F. Diederich
Laboratorium für Organische Chemie, ETH-Hönggerberg, HCI
8093 Zürich (Switzerland)
Fax: (+41) 163-21-109
E-mail: diederich@org.chem.ethz.ch

Supporting information for this article is available on the WWW under http://www.chembiochem.com or from the author.

  DOI: 10.1002/cbic.200300907

rings greatly affects aromatic–aromatic interactions by influencing the electronic nature of the rings.[14–16]

We recently started to systematically exchange one or more H atoms for fluorine in all possible positions of a previously developed family of thrombin inhibitors[17] in order to explore the fluorophilicity/fluorophobicity of the active site of this trypsin-like serine protease from the blood coagulation cascade (Figure 1). The rigid tricyclic core ensures that differently decorated inhibitors adopt nearly identical binding geometries in the structurally well-defined active site; a requirement for meaningful structure–activity relationships. Here, we report the results of the fluorine scan involving the benzylic residue undergoing edge-to-face interactions with the indole ring of Trp215 in the spacious lipophilic D pocket of the enzyme.[18] By using a combination of solution binding assays, X-ray crystal structure analysis of a protein–ligand complex, small-molecule X-ray crystallography, and data-base mining, we show that H–C_α–C=O fragments in proteins provide a pronounced fluorophilic environment. We provide evidence that C–F groups undergo favorable dipolar interactions with electrophilic residues such as C=O, C≡N, C–F, and even $-NO_2$, with the F atom pointing in a distinct orientation onto the positively polarized electrophilic center.

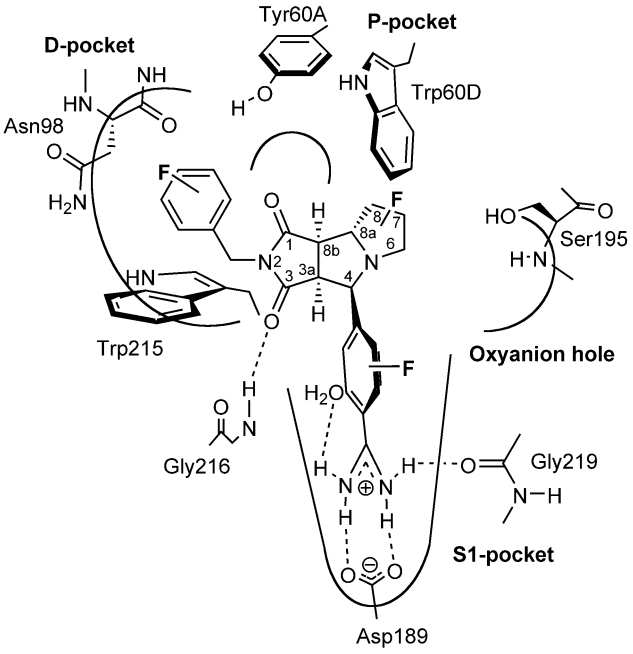
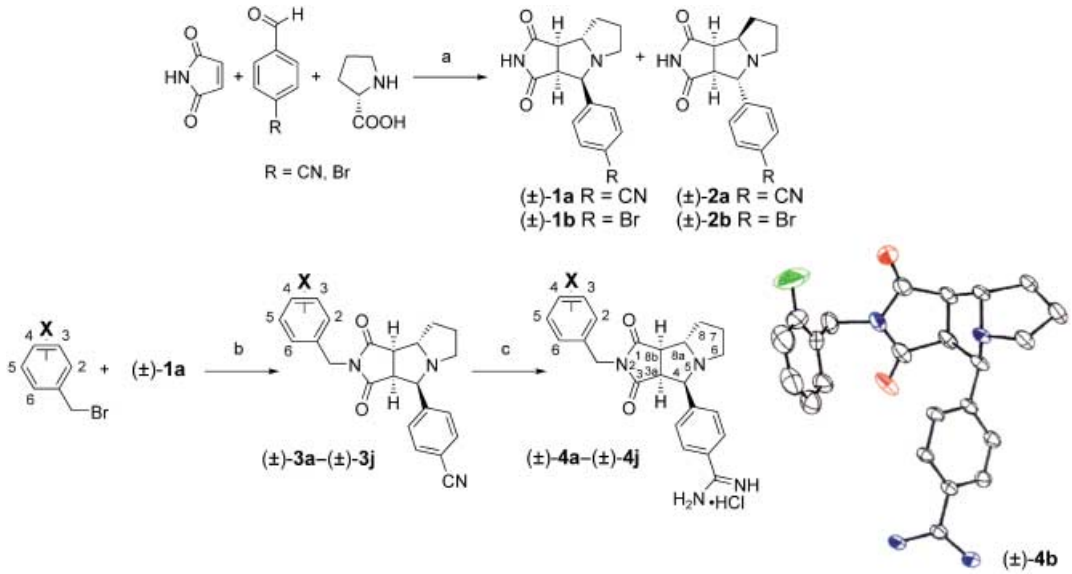


Figure 1. *Fluorine scan of tricyclic thrombin inhibitors and the schematic representation of the binding mode in the active site of thrombin. Only the (3aS,4R,8aS,8bR)-configured enantiomer is bound, according to X-ray crystallography.[17a,b] In addition to the catalytic triad flanked by the oxyanion hole, the active site can be described in terms of the selectivity pocket S1, a small proximal pocket (P1), and a large hydrophobic distal (D) pocket.*

Results and Discussion

Synthesis of the inhibitors

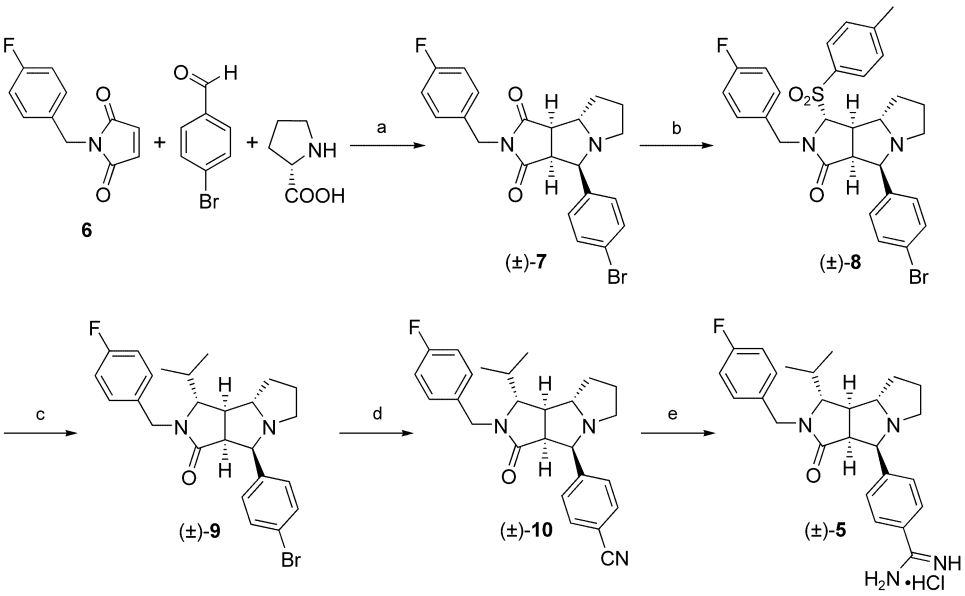
The fluorinated inhibitors were prepared from tricycle (±)-**1a**, which was obtained by 1,3-dipolar cycloaddition of maleimide, 4-formylbenzonitrile, and L-proline together with diastereoisomer (±)-**2a** (Scheme 1). The product ratio was solvent-dependent, varying from (±)-**1a**/(±)-**2a** 1:3 in MeCN to 1:4.3 in dimethylformamide (DMF). Following chromatographic separation, the assigned configuration of the diastereoisomers was confirmed by X-ray crystallography. A more favorable ratio (1:1.8) of the desired to undesired diastereoisomers (±)-**1b** and (±)-**2b** was obtained in MeCN, starting from *p*-bromobenzaldehyde. Interestingly, (±)-**1a** and (±)-**2a** have the same space group ($P2_1/c$) but very different crystal packings. The enantiomers of (±)-**1a** form dimers with two C=O···H–N hydro-



Scheme 1. *Synthesis of inhibitors (±)-**4a–j**. a) CH_3CN, reflux; 15% (±)-**1a** and 38% (±)-**2a**; b) K_2CO_3, [18]crown-6, toluene, reflux; 58–86%; c) i) AcCl, MeOH, CH_2Cl_2, 5°C; ii) NH_3 in MeOH, 65°C; 40–79%. Also shown is the ORTEP representation of (±)-**4b** with vibrational ellipsoids obtained at 203 K and shown at the 30% probability level. The two crystal bound water molecules have been omitted for clarity. For the substitution pattern X in (±)-**3a–j** and (±)-**4a–j**, see Table 1.*

gen bonds between their imide moieties (d(C=O···H–N) = 2.12 Å (150°)), while the enantiomers of (±)-**2a** are arranged in infinite chains, stabilized by C≡N···H–N hydrogen bonds (d(C≡N···H–N) = 2.16 Å (172°)) (see Supporting Information). Alkylation of (±)-**1a** with commercially available benzyl bromides (K_2CO_3, [18]crown-6, PhMe) afforded nitriles (±)-**3a–j** in high yields, and the Pinner reaction gave the desired phenylamidinium inhibitors, benzyl derivative (±)-**4a**,[17c] a series of eight mono-, di-, and pentafluorinated ligands (±)-**4b–i**, and monochlorinated (±)-**4j**. An X-ray crystal structure was obtained for the 2-fluorophenyl inhibitor (±)-**4b** and is shown in Scheme 1.

The synthesis of the tricyclic thrombin inhibitor (±)-**5**, which has an isopropyl group to fill the P-pocket[17b,c] started with the 1,3-dipolar cycloaddition of 4-fluorobenzylated maleimide **6**, 4-bromobenzaldehyde, and L-proline to give tricycle (±)-**7** (Scheme 2). Regioselective reduction with Li[Et_3BH], followed by treatment with 4-toluenesulfinic acid and $CaCl_2$ afforded sulfone (±)-**8**. Nucleophilic substitution (presumably via the acyliminium ion) with *i*-PrMgCl in the presence of $ZnCl_2$[19] provided (±)-**9**, which was converted into nitrile (±)-**10** and finally into the phenylamidinium salt (±)-**5**.



Scheme 2. *Synthesis of inhibitor (±)-**5**. a) CH_3CN, reflux; 38%; b) i) Li[Et_3BH], THF, −78→−45°C; ii) 4-toluenesulfinic acid, $CaCl_2$, CH_2Cl_2; 70% over 2 steps; c) iPrMgCl, $ZnCl_2$, CH_2Cl_2, 0°C; 71%; d) CuCN, DMF, reflux; 63%; e) AcCl, MeOH, CH_2Cl_2, 5°C; ii) NH_3 in MeOH, 65°C; 75%.*

Biological results and crystal structure of (±)-4d bound to the thrombin active site.

Inhibitory activity against thrombin and trypsin was determined as previously described by using the chromogenic substrate S-2238.[20] Surprisingly, whereas eight fluorinated and chlorinated inhibitors (inhibitory constant K_i = 0.19 to 0.61 μM) showed a potency similar to that of non-fluorinated (±)-**4a** (K_i = 0.31 μM), the 4-fluorophenyl derivative (±)-**4d** was much more active with K_i = 0.057 μM ($\Delta\Delta G_{(\pm)\text{-}4a \rightarrow (\pm)\text{-}4d} = -1.05 \pm 0.17$ kcal mol^{-1}) (Table 1). Furthermore, this compound showed the best selectivity (K_i(trypsin)/K_i(thrombin) = 67) against trypsin in the entire series. The introduction of the isopropyl residue in (±)-**5** to fill the P-pocket led to a large additional increase in affinity and selectivity, as expected from earlier work.[17] Ligand (±)-**5** with a K_i value of 5 nM and a 413-fold selectivity over trypsin is the most active in the series of tricyclic thrombin inhibitors prepared to date. Table 1 also includes measured log D values (logarithmic distribution coefficient at pH 7.4).[21] Due to the charged phenylamidinium residue, they are all in an unfavorable negative range. Since they do not vary much upon F substitution, the particularly increased binding affinity of (±)-**4d** does not result from F-mediated enhancement in lipophilicity. Also, possible differences in the edge-to-face interaction of the various fluorinated benzyl residues with the indole ring of Trp215 cannot explain the biological results.[17c,22,23]

Important insight into the origin of the enhanced affinity of (±)-**4d** was obtained by solving the X-ray crystal structure of the thrombin complex at 1.67 Å resolution.[24] Not surprisingly, only the (3a*S*,4-*R*,8a*S*,8b*R*) enantiomer is bound, as had already been shown for related thrombin inhibitors.[17a,b] Superimposition with a previously reported crystal structure of thrombin complexed with an inhibitor that has a piperonyl group pointing into the D pocket[17a] revealed an identical binding mode for both inhibitors (see Supporting Information). This corroborates our modeling-based assumption that all fluorinated inhibitors have identical binding geometries.

Table 1. Activities of the fluorinated thrombin inhibitors and selectivities with respect to trypsin. Also shown is the distribution coefficient LogD at pH 7.4, see[21]

Inhibitor	X	K_i [μM][a]	Selectivity[b]	Log D
(±)-**4a**	–	0.31	15	−1.24
(±)-**4b**	2-F	0.50	9.8	< −1
(±)-**4c**	3-F	0.36	26	−1.24
(±)-**4d**	4-F	0.057	67	−1.08
(±)-**4e**	2,3-F_2	0.49	18	n.d.[c]
(±)-**4f**	2,6-F_2	0.61	9.0	n.d.
(±)-**4g**	3,4-F_2	0.26	29	n.d.
(±)-**4h**	3,5-F_2	0.59	25	−1.25
(±)-**4i**	1,2,3,4,5-F_5	0.27	44	−1.14
(±)-**4j**	4-Cl	0.19	30	n.d.

[a] The uncertainty of measured K_i values is ±20%. [b] K_i(trypsin)/K_i(thrombin). [c] Not determined.

The F atom of (3a*S*,4*R*,8a*S*,8b*R*)-**4d** is in remarkably close contact with the H–C_α–C=O moiety of Asn98 (Figure 2), as indicated by the short distances d(F···C_α) = 3.1 Å (d(F···H–C_α) = 2.1 Å, α(F···H–C_α) = 157°) and d(F···C=O) = 3.5 Å (α(F···C=O) = 96°, α(F···C(O)–pseudotrigonal axis of the carbonyl system) = 28°).[25]

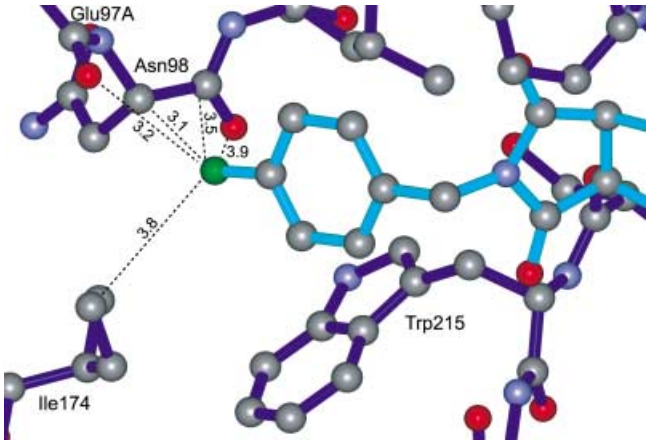


Figure 2. *Binding mode of the 4-fluorobenzyl moiety of (±)-**4d** in the D pocket of thrombin. All F contacts with the protein within a distance of < 4.0 Å are shown with dotted lines. Color code: dark-blue: thrombin skeleton, light-blue: inhibitor skeleton. gray: C atoms, blue: N atoms, red: O-atoms, green: F atom.*

What are the possible energetic contributions of these F contacts? The weak H-bond-donating ability of the polarized H–C_α unit in peptides is increasingly being recognized, and calculations indicate that weak X–O···H–C_α (X = H_2 and C) H-bonds may contribute roughly half of the strength of the conventional C=O···H–N hydrogen bond between two amide moieties.[26] Due to the very low polarizability of fluorine atoms, C–F residues are considerably weaker H-bond acceptors than C–O and C–N residues, and C–F···H–C is presumably better viewed as an attractive dipolar contact.[8] Nevertheless, C–F···H–C interactions have been reported to play an important role in crystal packing,[27] biological systems,[28] and catalysis.[29] Furthermore, intermolecular perturbation theory (IMPT) calculations in the Isostar Data Base suggest that interactions between the F atom of fluorobenzene and the H–C_α–C=O unit of acetone are energetically similar to the C–F···H–N interaction between fluorobenzene and *N*-methylacetamide.[30]

The C–F unit in (3a*S*,4*R*,8a*S*,8b*R*)-**4d** is also in close contact with the positively polarized C atom of the C=O unit of Asn98. Such organofluorine interactions with carbonyl groups have not previously been reported in the literature. However, a survey of the PDB revealed that such contacts are not uncommon.[31] Short C–F···C=O contacts, down to d(F···C) = 2.9 Å, are seen for both aromatic C–F residues as well as for $-CF_3$ groups. In several of these examples, the C–F bonds approach the carbonyl C atoms along the pseudotrigonal axis. Figure 3 shows the structure of an inhibitor of the trifluoroacetylpeptide class bound to porcine pancreatic eleastase at 2.50 Å resolution (PDB code: 2EST).[31c] All three F atoms of the CF_3CO group interact with backbone C=O groups of Cys191

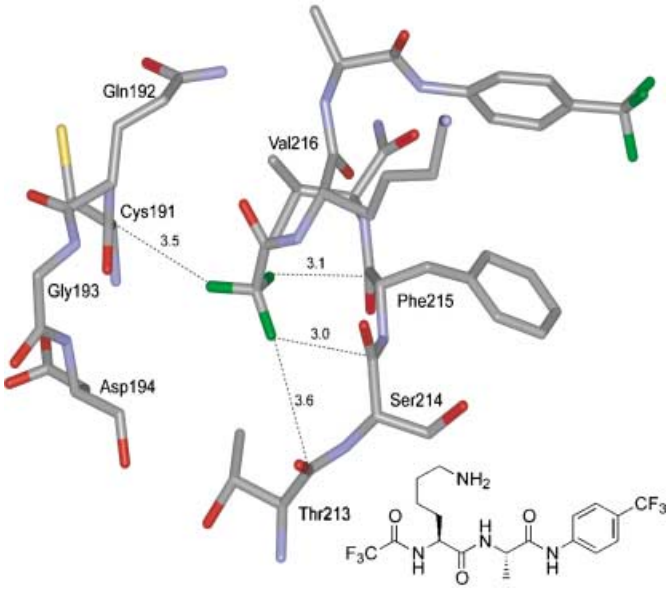


Figure 3. *C–F···C=O interactions seen in the crystal structure of the complex between porcine pancreatic elastase and an inhibitor of the trifluoroacetyl peptide class at 2.50 Å resolution (PDB code 2EST).[31c] Color code: gray: C atoms, blue: N atoms, red: O-atoms, green: F atom, yellow: S atom.*

(d(F···C) = 3.5 Å, α(F···C=O) = 59°), Thr213 (d(F···C) = 3.6 Å, α(F···C=O) = 87°), Ser214 (d(F···C) = 3.0 Å, α(F···C=O) = 74°), and Phe215 (d(F···C) = 3.1 Å, α(F···C=O) = 103°).

In view of these data, we suggest that the dipolar C–F···H–C_α and C–F···C=O interactions with the H–C_α–C=O moiety of Asn98 are the major determinant for the large increase in activity seen for (±)-**4d** when compared with the other inhibitors in the series. More structural information will be needed to explain why the ligands with different F-substitution patterns or additional F atoms (as in the 3,4-difluoro derivative **4g**), feature reduced binding affinities compared with (±)-**4d**.

C–F···C=O interactions in small-molecule X-ray crystal structures.

Additional support for the postulated C–F···H–C_α and C–F···C=O interactions was obtained by analyzing the X-ray crystal structures of the nitrile precursors (±)-**3d**, (±)-**3g**, (±)-**3i**, and (±)-**10**, inhibitor (±)-**4b**, and *N*-(4-fluorobenzyl)maleimide **6**.[32,33] In the crystal packing environments of (±)-**3d**, (±)-**3i**, and (±)-**10**, C–F···H–C_α–C=O and C–F···C=O contacts similar to the one seen in the thrombin complex of **4d** were found.[18] Only one close C–F···H–C_α–N contact below the sum of the van der Waals radii (2.67 Å) was observed in the crystal packing of compounds (±)-**3g** (d(F···H–C_α–N) = 2.52 Å, α(F···H–C–N) = 137.8°) and (±)-**4b** (d(F···H–C_α) = 2.29 Å, α(F···H–C_α) = 145.1°) (see Supporting Information).[25]

Each of the three molecules in the unit cell of *N*-(4-fluorobenzyl)maleimide **6** is engaged in intermolecular C–F···C=O contacts (d(F···C=O) = 2.94 Å, 3.11 Å, and 3.18 Å) (Figure 4).

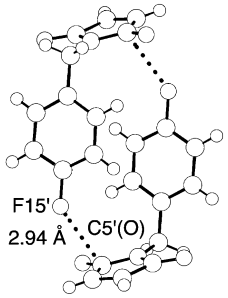


Figure 4. *Short intermolecular C-F···C=O contacts seen in the X-ray crystal packing of* N-*(4-fluorobenzyl)maleimide* ***6****.*

These contacts are significantly below the sum of the van der Waals radii (ca. 3.30 Å). Presumably they make a significant contribution to the energetics of the observed crystal packing.

A search in the CSD[34] for intermolecular C–F···C=O interactions in crystals of fluorine-containing organic carbonyl derivatives (ketones, esters, amides, urethanes, ureas, etc, as well as heterocyclic derivatives containing these structural elements) was carried out (see ref. [18] for a detailed discussion). A considerable number of C–F···C=O contacts (43 out of 1091 occurrences) with distances significantly below the sum of the van der Waals radii of C and F atoms (ca. 3.3 Å) were observed; this is indicative of an attractive interaction between the two polar groups. Both aliphatic and aromatic C–F units were found to be participating in these short contacts. The F···C=O interaction bears some structural resemblance to the interaction of nucleophiles with carbonyl groups as investigated by Bürgi and Dunitz,[35] In contrast to the latter, which involve partial transfer of electron density from the nucleophile to the electrophilic carbonyl system with concomitant rehybridization of the carbonyl group, the F···C=O contact is probably best described as a purely multipolar interaction between the intrinsically polar C–F and C=O units, without discernible structural changes in the carbonyl unit.

In our previous communication,[18] we used the scatterplot correlating the angle α_1(F···C=O) with the contact distance d_1(F···C(C=O)). While this scatterplot nicely shows a narrow cone of points, which suggests preferential positioning of the F atom near to 90° at close contacts, it does not discriminate between cases in which the F atom may approach the carbonyl group from the side, or even in the plane from those cases in which the F atom is truly located above or below the carbonyl unit at or near its pseudotrigonal axis. Therefore, we prefer the presentation shown in Figure 5a (left and center), here the distance d_2 of the F atom to the plane of the carbonyl

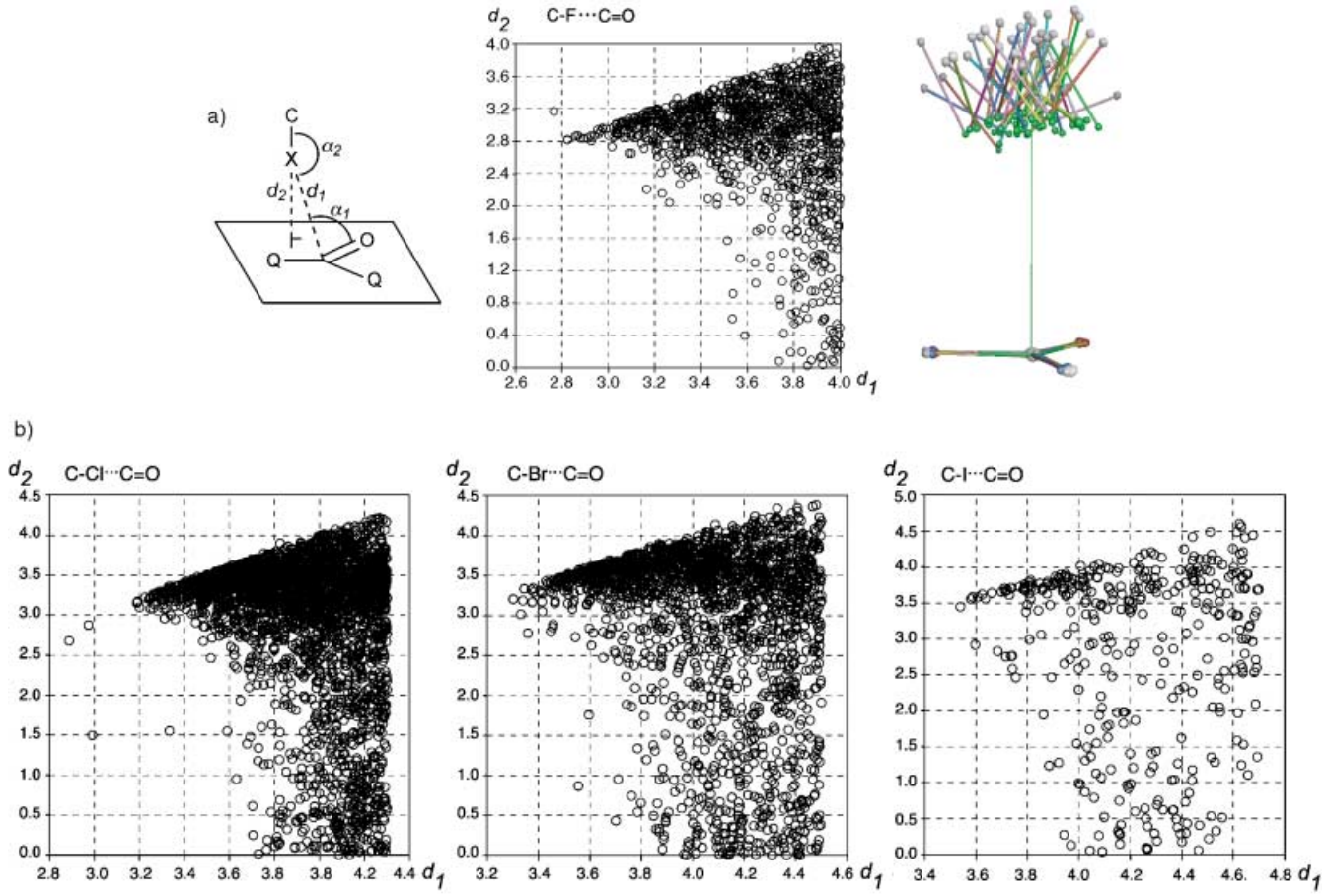


Figure 5. *A) left: Scatterplots of* d_2 *versus* d_1 *of close C–F···C=O contacts with* $2.5 < d_1 < 4.0$ *Å obtained from 628 hits of the CSD search of fluorine-containing organic carbonyl derivatives (Q=C, N, O). Right: Superposition of 43 intermolecular [C–F···Q(CO)Q] subunits with* $2.77 < d_1 < 3.09$ *Å extracted from the crystal structures and superimposed on the carbonyl unit.[18] The thin vertical line marks the pseudotrigonal axis of the carbonyl system (norm of the Q(CO)Q plane). B. Scatterplots of* d_2 *versus* d_1 *of close C–X···C=O (X=Cl, Br, I) contacts obtained from a CSD search of X-containing organic carbonyl derivatives (Q = C, N, O). C–Cl: 1701 hits (*$2.8 < d_1 < 4.3$ *Å) C–Br: 1148 hits (*$3.0 < d_1 < 4.5$ *Å). C–I: 193 hits (*$3.2 < d_1 < 4.7$ *Å).*

unit (defined by its three atoms Q–Q–O) is correlated with the internuclear distance d_1(F···C(CO). In this plot, a sharp tip results at close contact distances with a straight upper boundary for $d_2 \sim d_1$ for all those cases in which the F atom lies at or very close to the pseudotrigonal axis of the (planar) carbonyl unit. Points below this boundary belong to cases in which the F atom deviates from the pseudotrigonal axis ($d_2 < d_1$). In all these cases, the carbonyl unit adopts a nearly planar arrangement with the carbonyl C atom almost exactly in the plane of its three ligand atoms (distance from the plane 0–0.03 Å for all F···C contact distances). Very rarely, a point is found to lie above the linear boundary; this might occur when the carbonyl unit is significantly pyramidal, hence $d_2 > d_1$.

Superimposition of the crystal structures with short intermolecular contact distances of 2.9 Å $< d_1$(F···C=O) $<$ 3.1 Å with respect to the carbonyl system convincingly showed that the F atoms of the C–F unit prefer a position close to the pseudotrigonal axis. Figure 5a (right) shows the narrow cone around the pseudotrigonal axis of the carbonyl group (cone angle ca. 12°) within which the fluorine atom of a closely interacting C–F bond prefers to be located. Noticeably, at very short F···C contact distances, the C–F···C(O) angle α_2 lies between 110 and 150°. With increasing distance, this angle range widens, taking a value between 90 and 180°, near the van der Waals contact distance. The analysis also clearly revealed that a close contact of the F atom with the C=O group does not necessarily imply a close C–F···H–C_α contact.[18] Conversely, there are clear cases in which the F atom interacts predominantly with the C_α–H unit, but is somewhat more remote from the C=O group. Thus, there appears to be a range of possibilities for a C–F unit to engage in potentially attractive contacts with the H–C_α–C=O moiety.

It is interesting to compare the C–F···C=O case with those of C–Cl, C–Br, and C–I approaching a carbonyl system (Figure 5b). For all halogens, we can clearly identify the characteristic tip in the scatterplots that indicates those cases in which the C–X unit approaches the C=O group at short contact distances with X lying at or close to the pseudotrigonal axis of the carbonyl group. Note that the tip of the scatterplot occurs typically at distances significantly below the van der Waals contact distances; however, we note that, in going from F to the heavier halogens, the tip gets broader and the scatterplot widens more rapidly into a domain without preferential location of the X atom relative to the carbonyl unit ($0 < d_2 < d_1$). Interestingly, the onsets of these wide scatter domains occur at distances equal to or larger than the corresponding van der Waals contact distances for X···C. However, even at these larger distances some preference for C–X close to the pseudotrigonal axis of the carbonyl system remains, except in the case of C–I···C=O; here the scatterplot for $d_1 > 4.0$ Å becomes rather evenly distributed. The protrusion of the tip of the scatterplot from the onset of the wide, vertical (d_2) distribution is clearly most pronounced for the C–F case and is continuously reduced on going from C–F to C–Cl, C–Br, and C–I.

These results stimulated our interest in further investigating interactions between the C–F unit and electrophilic centers comparable to the C atom of the C=O unit.

Evidence for dipolar interactions between C–F units and other electrophilic centers from small-molecule X-ray crystallography

Intermolecular C–F···CN interactions, which strongly resemble the C–F···C=O interactions described above, were observed in the crystal-packing environment of several precursors to inhibitors, prepared as part of a fluorine scan to map the fluorophilicity/fluorophobicity of the S1-pocket in thrombin. An example is shown in Figure 6. Again, the F atom of (±)-**11** approaches the electrophilic C atom of the nitrile group in a neighboring

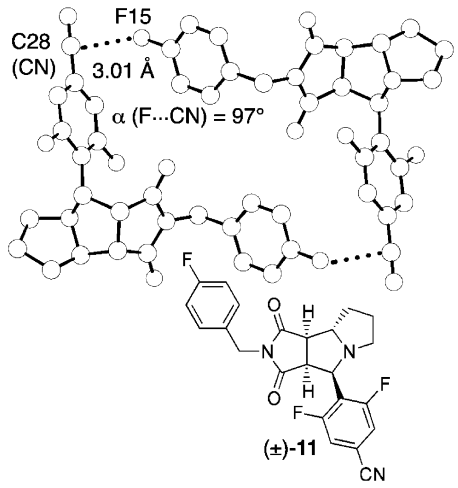


Figure 6. *C–F···CN interactions observed in the crystal packing environment of (±)-**11**.*

molecule at sub-van der Waals distance (d(F···C≡N = 3.01 Å) and in a nearly orthogonal fashion (α(F···C≡N) = 97°).[36] A similar contact had been noted by Nishide et al. as a novel intramolecular through-space interaction between F and CN in the X-ray crystallographic analysis of fluorocyanides.[37]

The C–F interactions observed in the X-ray crystal packing analyses of 14 fluorinated precursors prepared for the F scan at the thrombin active side and further supported by CSD searches can be divided into three main types (Figure 7). The first type comprises the interaction of C–F moieties with carbonyl and carboxyl derivatives as described above, as well as dipolar C–F···C–F interactions with similar geometrical preferences (Figure 7a). The second type comprises interactions between the F atom and positively polarized H atoms with a good number of contacts significantly below the sum of the van der Waals radii (2.67 Å) (Figure 7b). These contacts include C–F···H–C_α–C=O, C–F···H–C_α–N, and C–F···H–C_{Ar}, in particular if the positive polarization of the aromatic H atom is enhanced by a CN group in the *ortho*-position. In the third type, the positively polarized C atom of the C–F unit interacts with Cl^- ions or negatively polarized O atoms of carbonyl groups (Figure 7c). It is clear that further experimental and computational investigations are needed in order to determine the energetic contributions associated with these interactions and to explore their importance in chemical and biological systems.

Finally, we provide a first CSD search of the interaction of C–F units with organic nitro groups, which have an isosteric

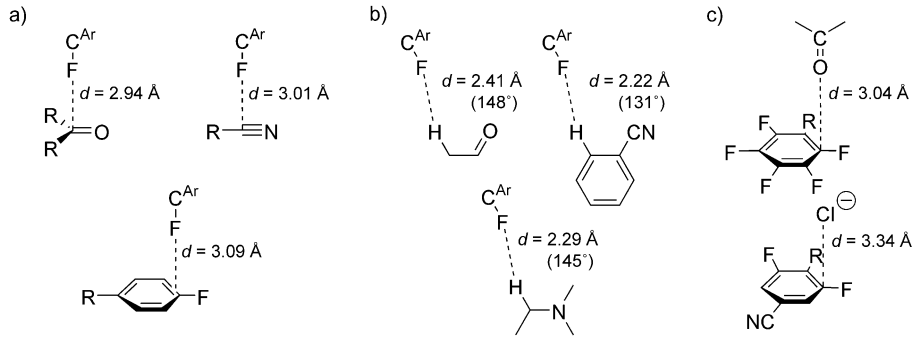


Figure 7. *Schematic representation of the three main types of intermolecular interactions observed for the C–F unit in the crystal packing environments of 14 different fluorinated synthetic precursors to tricyclic phenylamidinium inhibitors of thrombin. The shortest distances observed in the crystal structures are shown.*

resemblance to the carboxyl group. This analysis of intermolecular C–F···NO_2 contacts includes crystals of F-containing organic nitro derivatives with an intermolecular (nonbonded) distance between the C-bound F atom and the nitro N atom in the range 2.0 Å $< d$(F···NO_2) $<$ 4.0 Å.[34] According to our DFT calculations (RB3LYP, 6-31G** geometries),[38] the N atom in nitrobenzene carries a partial charge of +0.674, as derived from the molecular electrostatic potential, and a Mulliken charge of +0.387 and therefore possesses substantial electrophilic character.

The search resulted in 80 hits with 188 occurrences of F···NO_2 contacts and ten cases significantly below ($<$3.09 Å) the sum of the van der Waals radii of the N and F atoms (ca. 3.2 Å); this indicated an attractive interaction between the C–F unit and the NO_2 group (Figure 8). This dipolar interaction has, to the best of our knowledge, not been reported in the literature. The favored position of the F atom relative to the plane of the NO_2 group was further investigated. Comparing the scatterplot for C–X···NO_2 in Figure 8 with those of C–X···C=O (Figure 5), we note striking similarities. For C–F···NO_2, a relatively sharp tip protrudes from the "bulk" beginning at $d_1 \sim 3.4$ Å. However, the tip is not as sharp as in the case of the carbonyl system. Therefore, we may conclude that there is a marked tendency of the C–F unit to approach the electrophilic nitro group again more or less along the pseudotrigonal axis, but that this preference is somewhat less pronounced than in the case of a carbonyl unit.[39] In going to the heavier halogens, the scatterplots become much more diffuse, although a tip-forming cluster can always be recognized (not shown). Interestingly, for C–I···NO_2, a second cluster with $d_2 \sim 0$–0.5 Å is found at relatively short contact distances. This cluster relates to crystal structures in which the C–I unit approaches the nitro group in plane with short O···I contacts.[40] The closest C–F···NO_2 contact (2.67 Å) found in the search, appears from a visual inspection not to be controlled by other crystal packing forces.[41,42]

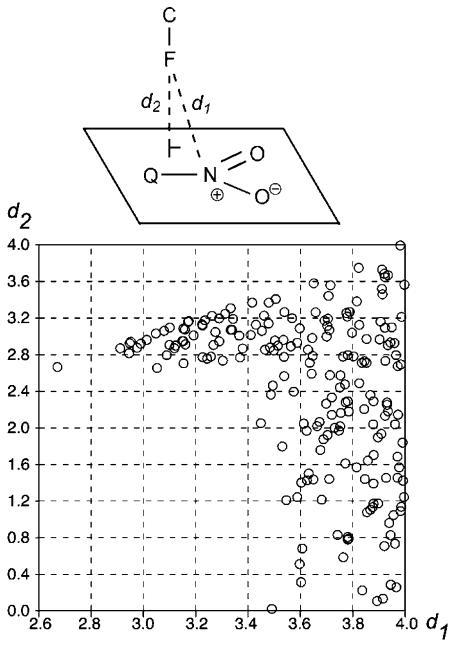


Figure 8. *Scatterplots of* d_2 *versus* d_1 *of close C–F···NO_2 contacts with* $2.0 < d_1 < 4.0$ *Å obtained from 86 hits of the CSD search of fluorine-containing organic nitro derivatives (Q = C).*

Conclusion

A fluorine scan of thrombin inhibitors to map favorable F environments in the D pocket of the enzyme has revealed that H–C_α–C=O fragments provide a pronounced fluorophilic environment. This finding, resulting from an X-ray crystal-structure analysis of a thrombin–ligand complex, was further corroborated by small-molecule X-ray crystallography and PDB and CSD searches. The previously unrecognized favorable C–F···C=O contacts displayed by a variety of carbonyl and carboxyl derivatives as well as by aliphatic and aromatic F substituents are best described in terms of multipolar interactions between the intrinsically polar C–F and C=O units. As the F atom approaches the carbonyl C atom, it is preferentially located near the pseudotrigonal axis of the carbonyl system. In consideration of the high propensity of H–C_α–C=O units in the active sites of proteins, such F interactions can be effectively exploited in medicinal chemistry for enhancing ligand affinity and/or selectivity in lead optimization.

Small-molecule X-ray crystallography and CSD searches show that the oriented interaction between aliphatic or aromatic C–F units and electrophilic centers below van der Waals contact distances is a more general one. F atoms prefer to approach C≡N units at an angle close to 90°. Favorable dipolar C–F···C–F contacts with similar orientation are observed, and the dipolar interaction between C–F units and NO_2 groups closely resembles, in its geometric preference, the one between C–F and C=O groups. On the other hand, Cl^- anions and carbonyl O atoms have been found to approach the electrophilic C atom of C–F units in an orthogonal alignment. The

study has further confirmed the propensity of C–F units to undergo dipolar interactions with H–C residues, contacts that are increasingly classified as weak H-bonds. A preference of C–F to undergo such interactions with aromatic C–H units that are particularly polarized by *ortho*-substituents such as CN, is clearly visible in the small-molecule crystal structures.

While it was important in a first step to identify the interaction preferences of C–F units, further studies must focus on the energetic quantification of the strength of these intermolecular contacts. For this purpose, we have recently initiated the design, synthesis, and study of specific model systems.

Experimental Section

General: Fluorinated benzyl bromides were purchased from Aldrich. Other solvents and reagents were of reagent grade, purchased from commercial suppliers and used without further purification. CH_2Cl_2 was freshly distilled from CaH_2, and toluene was distilled from Na. Anhydrous MeOH was purchased from Aldrich. Column chromatography (CC) on SiO_2–60 (230–400 mesh, 0.040–0.063 mm) from Fluka. TLC on SiO_2–60 F_{245}, Merck, visualization by UV light at 245 nm or by staining with 5% phosphomolybdic acid in EtOH. Melting points were determined on a Büchi SMP-20 and are uncorrected. In several cases, the compound decomposed before melting. IR spectra (cm^{-1}) were obtained with a Perkin–Elmer 1600-FTIR instrument. ^1H, ^{13}C, and ^{19}F NMR spectra were obtained on a Varian Gemini 300 with the solvent peak as internal reference, or $CFCl_3$ for the ^{19}F spectra. In the ^{13}C NMR spectra of (±)-**3e**, (±)-**3j**, (±)-**4b**, (±)-**4f**, (±)-**4g**, (±)-**4i**, and (±)-**8**, one aromatic signal is missing due to overlap, and only one of the signals of the F-bound C atoms in compound (±)-**4i** is reported. High-resolution MALDI mass spectra (HRMS) were obtained with 2,5-dihydroxybenzoic acid as matrix on a IonSpec Ultima instrument. Molecular ions (M^+) reported for phenylamidinium salts refer to the corresponding phenylamidine derivatives. Elemental analyses were performed by the Mikrolabor at the Laboratorium für Organische Chemie, ETH Zürich. All syntheses and characterizations of compounds not included in the manuscript are reported as Supporting Information.

(3a*SR*,4*RS*,8a*SR*,8b*RS*)-4-(1,3-Dioxodecahydropyrrolo[3,4-*a*]pyrrolizin-4-yl)benzonitrile ((±)-1a) and (3a*SR*,4*SR*,8a*RS*,8b*RS*)-4-(1,3-dioxodecahydropyrrolo[3,4-*a*]pyrrolizin-4-yl)benzonitrile ((±)-2a): A solution of maleimide (2.42 g, 25.0 mmol), 4-formylbenzonitrile (3.28 g, 25.0 mmol), and L-proline (2.88 g, 25.0 mmol) in MeCN (75 mL) was heated at reflux for 16 h. The solvent was removed in vacuo, and (±)-**1a** and (±)-**2a** were separated by CC (eluent EtOAc/Et_2O, 15:85).

(±)-**1a**: Yield: 1.05 g (15%), colorless crystals; m.p. >220°C (dec., EtOAc/Et_2O); ^1H NMR (300 MHz, $(CD_3)_2SO$): δ = 1.56–1.72 (m, 2H), 1.88–2.04 (m, 2H), 2.44–2.50 (m, 1H; H–C(6)), 2.70–2.82 (m, 1H; H–C(6)), 3.28–3.36 (m, 1H; H–C(8b)), 3.49–3.58 (m, 1H; H–C(8a)), 3.61 (t, *J* = 8.4, 1H; H–C(3a)), 4.17 (d, *J* = 8.7, 1H; H–C(4)), 7.51, 7.74 (AA′BB′, *J* = 8.1, 4H), 11.12 (s, 1H; NH); ^{13}C NMR (75 MHz, $(CD_3)_2SO$): δ = 24.2, 30.2, 50.9, 51.4, 52.6, 68.4, 68.5, 110.7, 120.0, 129.9, 132.6, 146.3, 177.8, 180.7; IR (KBr): 2228, 1770, 1706, 1607, 1367, 1349, 1206, 1189 cm^{-1}; HRMS: calcd for $C_{16}H_{16}N_3O_2$: 282.1243 $[MH]^+$; found 282.1239; X-ray.

(±)-**2a**: Yield: 2.69 g (38%), colorless crystals; m.p. >220°C (dec., EtOAc/Et_2O); ^1H NMR (300 MHz, $(CD_3)_2SO$): δ = 1.54–1.76 (m, 2H), 1.78–2.02 (m, 2H), 2.50–2.60 (m, 1H; H–C(6)), 2.78–2.98 (m, 1H; H–C(6)), 3.29 (dd, *J* = 9.0 and 6.9, 1H; H–C(3a)), 3.62 (t, *J* = 9.0, 1H; H–C(8b)), 3.68–3.78 (m, 1H; H–C(8a)), 4.09 (d, *J* = 6.9, 1H; H–C(4)), 7.63, 7.82 (AA′BB′, *J* = 8.1, 4H; Ar), 11.21 (s, 1H; NH); ^{13}C NMR (75 MHz, $(CD_3)_2SO$): δ = 25.4, 27.6, 49.7, 52.3, 57.2, 66.5, 69.2, 110.9, 119.8, 129.1, 133.3, 149.5, 179.2, 180.0; IR (KBr): 2228, 1771, 1704, 1604, 1363, 1342, 1193 cm^{-1}; HRMS: calcd for $C_{16}H_{16}N_3O_2$: 282.1243 $[MH]^+$; found 282.1236; elemental analysis calcd (%) for $C_{16}H_{15}N_3O_2$: C 68.31, H 5.37, N 14.94; found: C 68.46, H 5.48, N 14.80; X-ray.

(3a*SR*,4*RS*,8a*SR*,8b*RS*)-4-(2-(4-Fluorobenzyl)-1,3-dioxodecahydropyrrolo[3,4-*a*]pyrrolizin-4-yl)benzonitrile ((±)-3d): A mixture of (±)-**1a** (140.7 mg, 0.5 mmol), [18]crown-6 (13.2 mg, 0.05 mmol), K_2CO_3 (103.7 mg, 0.75 mmol), and 4-fluorobenzyl bromide (65 μL, 0.525 mmol) in toluene (2.5 mL) was stirred under N_2 at 100°C for 2 h. After addition of H_2O (5 mL) and EtOAc (5 mL) and phase separation, the aqueous phase was extracted with EtOAc (3×5 mL). The combined organic phases were dried (Na_2SO_4) and concentrated in vacuo, and recrystallization (EtOAc/hexane) afforded pure (±)-**3d** (165 mg, 85%) as colorless crystals. M.p. 164–165°C (EtOAc/hexane); ^1H NMR (300 MHz, $CDCl_3$): δ = 1.60–1.88 (m, 2H), 1.96–2.21 (m, 2H), 2.54–2.56 (m, 1H), 2.81–2.93 (m, 1H), 3.31 (d, *J* = 7.8, 1H), 3.52 (t, *J* = 8.4, 1H), 3.77 (dd, *J* = 10.4 and 7.1, 1H), 4.10 (d, *J* = 8.7, 1H), 4.46, 4.51 (AB, *J* = 13.8, 2H), 6.98 (t, *J* = 8.7, 2H; Ar), 7.22–7.28 (m, 2H), 7.30, 7.52 (AA′BB′, *J* = 8.1, 4H); ^{13}C NMR (75 MHz, $CDCl_3$): δ = 23.4, 29.6, 41.7, 49.0, 50.5, 50.8, 67.9, 68.3, 111.4, 115.3 (d, *J* = 21.4), 118.8, 128.6, 130.7 (d, *J* = 7.9), 131.3 (d, *J* = 3.6), 131.8, 143.5, 162.2 (d, *J* = 246.6), 174.6, 177.3; ^{19}F NMR (282 MHz, $CDCl_3$ + $CFCl_3$): δ = −113.5 (tt, *J* = 9.1 and 4.8); IR (KBr): 2224, 1774, 1704, 1604, 1510, 1399, 1341, 1297, 1225, 1169, 1160 cm^{-1}; HRMS: calcd for $C_{23}H_{21}FN_3O_2$: 390.1618 $[MH]^+$; found 390.1607; X-ray.[18]

(3a*SR*,4*RS*,8a*SR*,8b*RS*)-4-(2-(4-Fluorobenzyl)-1,3-dioxodecahydropyrrolo[3,4-*a*]pyrrolizin-4-yl)benzamidine hydrochloride ((±)-4d): CH_2Cl_2 (0.5 mL) and MeOH (0.5 mL) were added under N_2 to a 10 mL oven-dried flask, and the solution was cooled to 0°C. AcCl (0.5 mL) was added followed by (±)-**3d** (117 mg, 0.30 mmol) after 5 min. The mixture was allowed to stand for 36 h at 5°C. Addition of Et_2O (10 mL) gave a white precipitate that was isolated by filtration, washed with Et_2O (2×10 mL), and dried in vacuo. The white solid together with dry MeOH (0.5 mL) and NH_3 (2.0 M) in MeOH (0.5 mL, 1.0 mmol) was stirred for 3.5 h at 65°. After the mixture had cooled, the resulting NH_4Cl was precipitated by addition of acetone (10 mL) and removed by filtration. The solvent was evaporated in vacuo, and the residue dissolved in EtOH (0.5 mL). Addition of Et_2O (10 mL) led to the precipitation of (±)-**4d** (105 mg, 79%) as colorless solid. M.p >185°C (dec.); ^1H NMR (300 MHz, CD_3OD): δ = 1.72–1.90 (m, 2H), 2.00–2.18 (m, 2H), 2.56–2.66 (m, 1H), 2.80–2.91 (m, 1H), 3.45 (d, *J* = 8.1, 1H), 3.69–3.77 (m, 2H), 4.29 (d, *J* = 8.7, 1H), 4.45, 4.51 (AB, *J* = 14.3, 2H), 7.04 (t, *J* = 8.9, 2H), 7.33 (dd, *J* = 8.9 and 5.3, 2H), 7.44, 7.64 (AA′BB′, *J* = 8.6, 4H); ^{13}C NMR (75 MHz, CD_3OD): δ = 24.2, 30.5, 42.4, 50.3, 51.9, 52.0, 69.3, 69.5, 116.1 (d, *J* = 21.4), 128.3, 128.4, 130.2, 131.3 (d, *J* = 8.0), 133.3, 146.8, 163.5 (d, *J* = 244.1), 168.0, 177.0, 179.8; ^{19}F NMR (282 MHz, CD_3OD + $CFCl_3$): δ = 114.1 (tt, *J* = 8.5 and 5.4); IR (KBr): 3361, 3073, 1772, 1699, 1673, 1613, 1511, 1487, 1400, 1345, 1222, 1176 cm^{-1}; HRMS: calcd for $C_{23}H_{24}FN_4O_2$: 407.1883 $[MH]^+$; found 407.1875.

(3a*SR*,4*RS*,8a*SR*,8b*RS*)-4-(4-Bromophenyl)-2-(4-fluorobenzyl)-decahydropyrrolo[3,4-*a*]pyrrolizin-1,3-dione ((±)-7): A solution of **6** (3.08 g, 15.0 mmol), 4-bromobenzaldehyde (2.78 g, 15.0 mmol), and L-proline (1.73 g, 15.0 mmol) in MeCN (45 mL) was heated at reflux for 16 h. Concentration in vacuo and CC (CH_2Cl_2/EtOAc, 5:95→20:80) resulted in (±)-**7** (2.49 g, 38%) as colorless crystals. M.p. 154–156°C (CH_2Cl_2/EtOAc); ^1H NMR (300 MHz, $CDCl_3$): δ = 1.54–1.82 (m, 2H), 1.94–2.18 (m, 2H), 2.56–2.68 (m, 1H), 2.77–2.90

(m, 1H), 3.27 (dd, J=8.1 + 0.9, 1H), 3.47 (t, J=8.4, 1H), 3.75 (dd, J=9.9 and 7.2, 1H), 4.01 (d, J=8.7, 1H), 4.50 (s, 2H), 6.97 (dddd, J=9.0, 8.7, 3.0, and 2.4, 2H), 7.29 (ddd, J=8.7, 5.4, and 3.0, 2H), 7.06, 7.35 (AA′BB′, J=8.3, 4H); ^{13}C NMR (75 MHz, $CDCl_3$): δ=23.3, 29.4, 41.5, 48.9, 50.3, 50.6, 67.7, 68.0, 115.1 (t, J=21.2), 121.2, 129.5, 130.5 (d, J=8.5), 130.9, 131.3 (d, J=3.0), 136.8, 162.0 (d, J=244.4), 174.7, 177.5; ^{19}F NMR (282 MHz, $CDCl_3$ + $CFCl_3$): δ=−113.8 (*tt*, J=8.7 and 5.4); IR (KBr): 1770, 1708, 1605, 1510, 1485, 1428, 1397, 1337, 1222, 1170, 1102 cm^{-1}; HRMS: calcd for $C_{22}H_{20}BrFN_2O_3$: 443.0770 and 445.0750 [MH]$^+$; found 443.0769 and 445.0760; elemental analysis calcd (%) for $C_{22}H_{20}BrFN_2O_2$: C 59.61, H 4.55, N 6.32, F 4.29, Br 18.02; found: C 59.76, H 4.60, N 6.52, F 4.17, Br 18.19.

(1*RS*,3 a*SR*,4*RS*,8 a*SR*,8 b*RS*)-4-(4-Bromophenyl)-2-(4-fluorophenylmethyl)-1-[(4-methylphenyl)sulfonyl]-octahydropyrrolo[3,4-*a*]pyrrolizin-3-one ((±)-8). A solution of Li[Et_3BH] in THF (1.0 M, 9.5 mL, 9.5 mmol) was added gradually over 15 min to (±)-**7** (2.21 g, 5 mmol) in THF (20 mL) at −78 °C. After being stirred for 1 h, the mixture was allowed to warm to −45 °C and then quenched by addition of saturated aqueous $NaHCO_3$ (20 mL). The phases were separated, and the aqueous phase was further extracted with CH_2Cl_2 (3×20 mL). The combined organic phases were dried ($MgSO_4$) and concentrated in vacuo. 4-Toluenesulfinic acid (2.34 g, 15 mmol), dry powdered $CaCl_2$ (1.66 g, 15 mmol), and CH_2Cl_2 (30 mL) were added to the crude material. After the mixture had been stirred for 36 h under Ar at RT, $NaHCO_3$ (100 mL) and CH_2Cl_2 (60 mL) were added, and the mixture was filtered. The organic layer was separated, and the aqueous phase was extracted with CH_2Cl_2 (60 mL). The combined organic layers were washed with saturated aqueous $NaHCO_3$ (2×100 mL), dried (Na_2SO_4), concentrated in vacuo, and purified by CC (EtOAc/cyclohexane 30:70) to give (±)-**8** (2.03 g, 70%) as a colorless solid. M.p. 179–181 °C (EtOAc/cyclohexane); ^{1}H NMR (300 MHz, $CDCl_3$): δ=1.58–1.74 (m, 2H), 1.80–2.04 (m, 2H), 2.42–2.58 (m, 2H), 2.48 (s, 3H), 2.80–2.91 (m, 1H), 2.97 (dd, J=7.8 and 3.3, 1H), 2.98–3.07 (m, 1H), 3.90 (d, J=6.6, 1H), 4.27 (s, 1H), 4.13, 5.09 (AB, J=14.7, 2H), 6.91–7.09 (m, 4H), 7.15, 7.39 (AA′BB′, J=8.1, 4H), 7.40, 7.71 (AA′BB′, J=8.6, 4H); ^{13}C NMR (75 MHz, $CDCl_3$): δ=21.9, 24.5, 31.8, 43.0, 44.5, 50.9, 51.8, 69.5, 71.3, 80.9, 115.6 (t, J=21.9), 121.0, 129.2, 129.5, 130.0 (d, J=7.9), 130.3, 130.7, 131.9, 136.9, 146.0, 162.2 (d, J=244.5), 172.3; ^{19}F NMR (282 MHz, $CDCl_3$ + $CFCl_3$): δ=−113.8 (tt, J=8.5 and 5.4); IR (KBr): 1710, 1605, 1598, 1509, 1486, 1453, 1389, 1371, 1316, 1303, 1292, 1227, 1200, 1157, 1138 cm^{-1}; HRMS: calcd for $C_{29}H_{30}BrFN_2O_3S$: 583.1066 and 585.1046 [MH]$^+$; found 583.1068 and 585.1036.

(1*RS*,3 a*SR*,4*RS*,8 a*SR*,8 b*RS*)-4-(4-Bromophenyl)-2-(4-fluorophenylmethyl)-1-isopropyl-octahydropyrrolo[3,4-*a*]pyrrolizin-3-one ((±)-9. *i*PrMgCl in Et_2O (2.0 M, 1.5 mL, 3.0 mmol) was added dropwise to $ZnCl_2$ in Et_2O (1.0 M, 1.65 mL, 1.65 mmol) and CH_2Cl_2 (5.5 mL) in a dry flask under Ar, and a milky solution was formed. After 30 min, (±)-**8** (0.88 g, 1.5 mmol) in CH_2Cl_2 (5.5 mL) was added dropwise over 15 min at 0 °C. The mixture was stirred for 30 min at 0 °C and for 16 h at RT, then quenched by addition of HCl (1 M) until pH<1. After being stirred for 15 min, the acidic solution was neutralized with saturated aqueous $NaHCO_3$ solution. The phases were separated, and the aqueous phase was further extracted with CH_2Cl_2 (3×40 mL). The combined organic phases were dried (Na_2SO_4), concentrated in vacuo, and purified by CC (EtOAc/cyclohexane 30:70→40:60) to give (±)-**9** (0.50 g, 71%) as colorless crystals. M.p. 199–200 °C (EtOAc/cyclohexane); ^{1}H NMR (300 MHz, $CDCl_3$): δ=0.70 (d, J=6.9, 3H), 0.91 (d, J=6.9, 3H), 1.53–1.78 (m, 2H), 1.88–2.11 (m, 3H), 2.48 (dt, J=8.7 + 2.7, 1H), 2.57–2.67 (m, 1H), 2.87–2.98 (m, 1H), 3.18–3.27 (m, 2H), 3.30 (t, J=8.1, 1H), 4.08 (d, J=7.8, 1H), 3.79, 4.82 (AB, J=15.2, 2H), 6.96–7.04 (m, 2H), 7.08–7.15 (m, 2H), 7.28, 7.44 (AA′BB′, J=8.4, 4H); ^{13}C NMR (75 MHz, $CDCl_3$): δ=14.7, 18.5, 24.6, 28.1, 31.4, 41.5, 43.3, 52.4, 52.7, 67.4, 70.0, 73.2, 115.2 (t, J=21.3), 120.6, 129.5 (d, J=8.6), 129.7, 130.7, 132.2 (d, J=3.1), 138.5, 161.9 (d, J=241.9), 172.3; ^{19}F NMR (282 MHz, $CDCl_3$ + $CFCl_3$): δ=−114.9 (tt, J=8.5 and 5.4); IR (KBr): 1673, 1508, 1484, 1446, 1432, 1418, 1351, 1252, 1214, 1157 cm^{-1}; HRMS: calcd for $C_{25}H_{29}BrFN_2O$, 471.1447 and 473.1427 [MH]$^+$; found 471.1454 and 473.1427; elemental analysis calcd (%) for $C_{25}H_{28}BrFN_2O$: C 63.70, H 5.90, Br 16.95, F 4.03, N 5.94; found: C 63.90, H 6.19, Br 16.86, F 4.02, N 6.17.

(1*RS*,3 a*SR*,4*RS*,8 a*SR*,8 b*RS*)-4-{2-(4-Fluorophenylmethyl)-3-oxo-1-isopropyloctahydropyrrolo[3,4-*a*]pyrrolizin-4-yl}benzonitrile ((±)-10): A solution of (±)-**9** (0.41 g, 0.89 mmol) and CuCN (0.36 g, 1.78 mmol) in dry DMF (5 mL) was heated at reflux (bath 155 °C) under Ar for 18 h. The mixture was cooled to RT, and CH_2Cl_2 (6 mL) and conc. NH_3 (6 mL) were added. After the mixture had been stirred for 30 min, additional CH_2Cl_2 (30 mL), conc. NH_3 (10 mL), and saturated aqueous NaCl (20 mL) were added. The phases were separated, and the aqueous phase was further extracted with CH_2Cl_2 (2×30 mL). The combined organic phases were washed with conc. NH_3 (50 mL) and saturated aqueous NaCl (2×50 mL), dried (Na_2SO_4), and concentrated in vacuo. CC (EtOAc/cyclohexane 40:60→50:50) resulted in (±)-**10** (235 mg, 63%) as light yellow crystals. M.p. 148–150 °C (EtOAc/cyclohexane); ^{1}H NMR (300 MHz, $CDCl_3$): δ=0.80 (d, J=6.9, 3H), 1.02 (d, J=6.9, 3H), 1.60–1.84 (m, 2H), 1.98–2.22 (m, 3H), 2.55–2.63 (m, 1H), 2.63–2.74 (m, 1H), 3.00–3.11 (m, 1H), 3.28–3.38 (m, 2H), 3.44 (t, J=8.1, 1H), 4.25 (d, J=7.5, 1H), 3.90, 4.90 (AB, J=15.2, 2H), 6.96–7.16 (m, 2H), 7.20–7.30 (m, 2H), 7.62, 7.71 (AA′BB′, J=8.1, 4H); ^{13}C NMR (75 MHz, $CDCl_3$): δ=14.8, 18.5, 24.7, 28.1, 31.4, 41.3, 42.3, 52.4, 52.9, 67.4, 70.3, 73.4, 110.4, 115.3 (t, J=21.2), 119.2, 128.6, 129.4 (d, J=7.9), 131.4, 132.1 (d, J=3.1), 145.5, 161.8 (d, J=243.2), 172.0; ^{19}F NMR (282 MHz, $CDCl_3$ + $CFCl_3$): δ=−114.7 (tt, J=8.5 and 5.4); IR (KBr): 2225, 1674, 1608, 1601, 1509, 1446, 1432, 1418, 1254, 1215, 1158, 1125 cm^{-1}; HRMS: calcd for $C_{26}H_{29}FN_3O$: 418.2294 [MH]$^+$; found 418.2294; elemental analysis calcd (%) for $C_{26}H_{28}FN_3O$: C 74.79, H 6.76, F 4.55, N 10.06; found: C 74.75, H 6.80, F 4.44, N 9.91; X-ray.[18]

(1*RS*,3 a*SR*,4*RS*,8 a*SR*,8 b*RS*)-4-{2-(4-Fluorophenylmethyl)-3-oxo-1-isopropyloctahydropyrrolo[3,4-*a*]pyrrolizin-4-yl}benzamidine hydrochloride ((±)-5): Inhibitor (±)-**5** (125 mg and 0.30 mmol) was synthesized from (±)-**10** according to the procedure described for (±)-**4d**. Yield: 105 mg (75%), colorless solid; m.p. >150 °C (dec.); ^{1}H NMR (300 MHz, CD_3OD): δ=0.74 (d, J=6.9, 3H), 0.97 (d, J=6.9, 3H), 1.60–1.84 (m, 2H), 1.96–2.10 (m, 2H), 2.12–2.24 (m, 1H), 2.56–2.67 (m, 1H), 2.69 (dt, J=8.4 and 2.5, 1H), 2.89–3.00 (m, 1H), 3.24–3.38 (m, 2H), 3.48 (t, J=8.1, 1H), 4.32 (d, J=7.5, 1H), 3.97, 4.71 (AB, J=15.2, 2H), 7.04–7.14 (m, 2H), 7.26–7.35 (m, 2H), 7.64, 7.74 (AA′BB′, J=8.3, 4H); ^{13}C NMR (75 MHz, CD_3OD): δ=15.0, 18.6, 25.4, 29.3, 32.0, 42.3, 44.2, 53.4, 54.7, 69.6, 71.5, 75.1, 116.4 (d, J=21.8), 127.8, 128.3, 130.4, 131.1 (d, J=7.9), 133.7 (d, J=3.1), 148.4, 163.7 (d, J=243.2), 168.4, 174.9; ^{19}F NMR (282 MHz, CD_3OD + $CFCl_3$): δ=−114.5 (tt, J=8.5 and 5.4); IR (KBr): 3108, 1665, 1611, 1534, 1510, 1485, 1464, 1447, 1414, 1391, 1297, 1158 cm^{-1}; HRMS: calcd for $C_{26}H_{32}FN_4O$, 435.2560 [MH]$^+$; found 435.2561.

Acknowledgements

We thank Dr. B. Schweizer (ETH Zürich) for his help with the CSD searches and Dr. B. Kuhn (Roche) for assistance with the calculations of partial charges. This work was supported by the ETH Re-

search Council, F. Hoffmann–La Roche Ltd, and by the Carlsberg Foundation (postdoctoral fellowship for J.A.O.).

Keywords: enzyme inhibitors · fluorine · medicinal chemistry · molecular recognition · thrombin

[1] Derwent Drug File (Derwent subscriber file, DRUGU) covers the pharmaceutical literature going back to 1983, Derwent Information Limited, Derwent House, 14 Great Queen Street, London, WC2B 5DF (UK) (30.05.2003).
[2] a) K. Park, N. R. Kitteringham, P. M. O'Neill, *Annu. Rev. Pharmacol. Toxicol.* **2001**, *41*, 443–470; b) D. O'Hagan, H. S. Rzepa, *Chem. Commun.* **1997**, 645–652; c) B. E. Smart, *J. Fluorine Chem.* **2001**, *109*, 3–11.
[3] M. Schlosser, D. Michel, *Tetrahedron* **1996**, *52*, 99–108.
[4] F. T. T. Huque, K. Jones, R. A. Saunders, J. A. Platts, *J. Fluorine Chem.* **2002**, *115*, 119–128.
[5] M. R. Groves, Z.-J. Yao, P. P. Roller, T. R. Burke, Jr., D. Barford, *Biochemistry* **1998**, *37*, 17773–17783.
[6] A. Kashima, Y. Inoue, S. Sugio, I. Maeda, T. Nose, Y. Shimohigashi, *Eur. J. Biochem.* **1998**, *255*, 12–23.
[7] T. R. Burke, Jr., B. Ye, X. Yan, S. Wang, Z. Jia, L. Chen, Z.-Y. Zhang, D. Barford, *Biochemistry* **1996**, *35*, 15989–15996.
[8] J. D. Dunitz, R. Taylor, *Chem. Eur. J.* **1997**, *3*, 89–98.
[9] G. M. Dubowchik, V. M. Vrudhula, B. Dasgupta, J. Ditta, T. Chen, S. Sheriff, K. Sipman, M. Witmer, J. Tredup, D. M. Vyas, T. A. Verdoorn, S. Bollini, A. Vinitsky, *Org. Lett.* **2001**, *3*, 3987–3990.
[10] A. DerHovanessian, J. B. Doyon, A. Jain, P. R. Rablen, A.-M. Sapse, *Org. Lett.* **1999**, *1*, 1359–1362.
[11] J. A. Ippolito, D. W. Christianson, *Int. J. Biol. Macromol.* **1992**, *14*, 193–197.
[12] A. Jain, G. M. Whitesides, R. S. Alexander, D. W. Christianson, *J. Med. Chem.* **1994**, *37*, 2100–2105.
[13] J. Ren, J. Milton, K. L. Weaver, S. A. Short, D. I. Stuart, D. K. Stammers, *Structure* **2000**, *8*, 1089–1094.
[14] a) C.-Y. Kim, J. S. Chang, J. B. Doyon, T. T. Baird, Jr., C. A. Fierke, A. Jain, D. W. Christianson, *J. Am. Chem. Soc.* **2000**, *122*, 12125–12134; b) R. D. Madder, C.-Y. Kim, P. P. Chandra, J. B. Doyon, T. A. Baird, Jr., C. A. Fierke, D. W. Christianson, J. G. Voet, A. Jain, *J. Org. Chem.* **2002**, *67*, 582–584.
[15] B. C. Finzel, E. T. Baldwin, G. L. Bryant Jr. , G. F. Hess, J. W. Wilks, C. M. Trepod, J. E. Mott, V. P. Marshall, G. L. Petzold, R. A. Poorman, T. J. O'Sullivan, H. J. Schostarez, M. A. Mitchell, *Protein Sci.* **1998**, *7*, 2118–2126.
[16] For a review, see: E. A. Meyer, R. K. Castellano, F. Diederich, *Angew. Chem.* **2003**, *115*, 1244–1287; *Angew. Chem. Int. Ed.* **2003**, *42*, 1210–1250.
[17] a) U. Obst, V. Gramlich, F. Diederich, L. Weber, D. W. Banner, *Angew. Chem.* **1995**, *107*, 1874–1877; *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 1739–1742; b) U. Obst, D. W. Banner, L. Weber, F. Diederich, *Chem. Biol.* **1997**, *4*, 287–295; c) U. Obst, P. Betschmann, C. Lerner, P. Seiler, F. Diederich, V. Gramlich, L. Weber, D. W. Banner, P. Schönholzer, *Helv. Chim. Acta* **2000**, *83*, 855–909; d) P. Betschmann, S. Sahli, F. Diederich, U. Obst, V. Gramlich, *Helv. Chim. Acta* **2002**, *85*, 1210–1245.
[18] For a preliminary communication of parts of this work, see: J. A. Olsen, D. W. Banner, P. Seiler, U. Obst Sander, A. D'Arcy, M. Stihle, K. Müller, F. Diederich, *Angew. Chem.* **2003**, *115*, 2611–2615; *Angew. Chem. Int. Ed.* **2003**, *42*, 2507–2511.
[19] D. S. Brown, P. Charreau, T. Hansson, S. V. Ley, *Tetrahedron* **1991**, *47*, 1311–1328.
[20] K. Hilpert, J. Ackermann, D. W. Banner, A. Gast, K. Gubernator, P. Hadváry, L. Labler, K. Müller, G. Schmid, T. B. Tschopp, H. van de Waterbeemd, *J. Med. Chem.* **1994**, *37*, 3889–3901.
[21] $\log D_{pH} = \log P - \log[1+10^{(pH-pK_a)}]$. $\log P$ = octanol/water partition coefficient. $\log D$ provides an indicator of the lipophilicity of a drug at the pH of blood plasma.
[22] For changes in atomic charge densities in Phe as a result of fluorination, see: A. Matsushima, T. Fujita, T. Nose, Y. Shimohigashi, *J. Biochem.* **2000**, *128*, 225–232.
[23] S. Paliwal, S. Geib, C. S. Wilcox, *J. Am. Chem. Soc.* **1994**, *116*, 4497–4498.
[24] The experimental details are found in ref. [18]. PDB code: 1OYT.
[25] a) P. R. Gerber, K. Müller, *J. Comput.-Aided Des.* **1995**, *9*, 251–268; Gerber Molecular Design (http://www.moloc.ch). The distances and angles were obtained with the Moloc modeling/visualization program. The position of H_α of Asn98 is based on stereochemical consideration with a C–H bond length of 1.08 Å.
[26] a) T. Steiner, *Acta Crystallogr. D* **1995**, *51*, 93–97; b) R. Vargas, J. Garza, D. A. Dixon, B. P. Hay, *J. Am. Chem. Soc.* **2000**, *122*, 4750–4755; c) S. Scheiner, T. Kar, Y. Gu, *J. Biol. Chem.* **2001**, *276*, 9832–9837.
[27] V. R. Thalladi, H.-C. Weiss, D. Bläser, R. Boese, A. Nangia, G. R. Desiraju, *J. Am. Chem. Soc.* **1998**, *120*, 8702–8710.
[28] J. Parsch, J. W. Engels, *J. Am. Chem. Soc.* **2002**, *124*, 5664–5672.
[29] a) V. V. Grushin, W. J. Marshall, *Angew. Chem.* **2002**, *114*, 4656–4659; *Angew. Chem. Int. Ed.* **2002**, *41*, 4476–4479; b) A. J. Mountford, D. L. Hughes, S. J. Lancaster, *Chem. Commun.* **2003**, 2148–2149.
[30] Isostar Database version 1.6, November 2003, Cambridge Crystallographic Data Centre, Union Road 12, Cambridge, CBZ 1EZ (UK).
[31] For example, see: a) M. Adler, D. D. Davey, G. B. Philips, S.-H. Kim, J. Jancarik, G. Rumennik, D. R. Light, M. Whitlow, *Biochemistry* **2000**, *39*, 12534–12542, PDB code 1FJS; b) Z. Wang, B. J. Canagarajah, J. C. Boehm, S. Kassisà, M. H. Cobb, P. R. Young, S. Abdel-Meguid, J. L. Adams, E. J. Goldsmith, *Structure* **1998**, *6*, 1117–1128, PDB code 1BL7; c) D. H. Hughes, L. C. Sieker, J. Bieth, J.-L. Dimicoli, *J. Mol. Biol.* **1982**, *162*, 645–658, PDB code 2EST; d) C. Mattos, G. A. Petsko, D. Ringe, *Biochemistry* **1995**, *34*, 3193–3203, PDB code 1ELE; e) PDB searches were conducted by using Relibase+, V.1.2.3., January 2004, CCSD, Cambridge, **2004**.
[32] The crystal structures of compounds (±)-**3d**, (±)-**3i**, and (±)-**10** have been reported previously.[18] CCDC-226080 ((±)-**1a**), 226079 ((±)-**2a**), 226082 ((±)-**3g**), 226083 ((±)-**4b**), and 226082 (**6**) contain the supplementary crystallographic data for this paper. These data can be obtained free of charge via www.ccdc.cam.ac.uk/conts/retrieving.html (or from the Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge CB2 1EZ, UK; fax: (+44) 1223-336033; or deposit@ccdc.cam.uk). Further details are found as Supporting Information.
[33] For an earlier crystallographic investigation, see: S. Larsen, K. Marthi, *Acta Crystallogr. B* **1994**, *50*, 373–381.
[34] CSD Version 5.24 of November 2002, holding over 270000 small-molecule crystal structures. The search was limited to well-resolved (*R* factor $<10\%$, no disorder) neutral organic molecules, excluding metalorganic or polymeric compound.
[35] B. Bürgi, J. D. Dunitz, *Acc. Chem. Res.* **1983**, *16*, 153–161.
[36] For a description of the fluorine scan of the phenylamidinium residue occupying the S1-pocket of thrombin, containing the detailed description of the X-ray crystal structure of (±)-**11**, see: J. Olsen, P. Seiler, B. Wagner, T. Tschopp, H. Fischer, U. Obst Sander, D. W. Banner, M. Kansy, K. Müller, F. Diederich, *Org. Biomol. Chem.*, **2004**, in press.
[37] K. Nishide, Y. Hagimoto, H. Hasegawa, M. Shiro, M. Node, *Chem. Commun.* **2001**, 2394–2395.
[38] Spartan '04 Windows, **2003**, Wavefunction, Inc., Irvine, CA (USA).
[39] It was pointed out to us by one referee that strong librational motions of the (always) terminal NO_2 groups (as compared to peptide bonds that are often involved in strong additional interactions with neighboring molecules) could result in the more diffuse scatterplots.
[40] For examples of in-plane C–I···O_2N– contacts with short I···O distances, see: G. R. Desiraju, *Angew. Chem.* **1995**, *107*, 2541–2558; *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 2311–2327.
[41] Nucleophilic oxygen has previously been shown to approach the N atom of nitro groups at sub-van der Waals distance in a fashion similar to that described here for the C–F···NO_2 contacts; see: J. A. Platts, S. T. Howard, K. Wozniak, *Chem. Phys. Lett.* **1995**, *232*, 479–485.
[42] It has been observed by X-ray crystallography of carbocations with SbF_6^- counteranions that F atoms interact with the cationic center in an orthogonal way, see: R. F. Childs, M. D. Kostyk, C. J. L. Lock, M. Mahendran, *J. Am. Chem. Soc.* **1990**, *112*, 8912–8920.

Received: December 15, 2003 [F907]

