Mapping Enzyme Active Sites

A Fluorine Scan of Thrombin Inhibitors to Map the Fluorophilicity/Fluorophobicity of an Enzyme Active Site: Evidence for C–F···C=O Interactions**

Jacob A. Olsen, David W. Banner, Paul Seiler, Ulrike Obst Sander, Allan D'Arcy, Martine Stihle, Klaus Müller,* and François Diederich**

The modulation of pharmacokinetic properties by fluorine substitution has become a well-established strategy for lead optimization[1] and has resulted in a large number (ca. 150) of fluorinated drugs in clinical use. The substitution of hydrogen by fluorine often causes minimal steric effects despite their different van der Waals radii (1.20 and 1.47 Å, respectively),[1,2] while a number of properties, such as lipophilicity, pK_a values, and metabolic stability, can be affected in a favorable way. Multiple fluorination of aromatic rings changes the quadrupole moment,[3] which profoundly affects aromatic–aromatic interactions.[4] The bioisosteric replacement of hydrogen by fluorine also affects enzyme–ligand binding affinity and selectivity, as shown for inhibitors of carboxypeptidase A,[5] stromelysin A,[6] carbonic anhydrase,[7] and chymotrypsin.[8]

In the same way as an active-site environment can be described as being hydrophilic or lipophilic, it could also be classified as being either fluorophilic or fluorophobic because of the unique properties of fluoro-organic compounds. Teflon and other fluorinated compounds clearly are neither hydrophilic nor hydrophobic.[9] Understanding the nature and intrinsic properties of fluorophilic/fluorophobic environments is therefore of great interest for structure-based lead design and optimization.

We recently reported a family of inhibitors of thrombin with a rigid tricyclic core that show nearly identical binding geometries in the structurally well-defined active site of this trypsin-like serine protease from the blood coagulation cascade.[10,11] We have recently started to systematically exchange one or more H atoms for F atoms in various positions of these inhibitors to explore the fluorophilicity/fluorophobicity of the thrombin active site. Here we report the results of a fluorine scan involving the benzylic residue that undergoes edge-to-face interactions with the indole ring of Trp 215 in the spacious hydrophobic D pocket of the enzyme (Figure 1).[12]

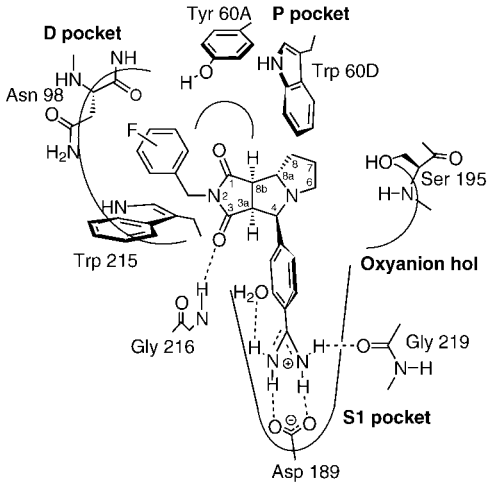
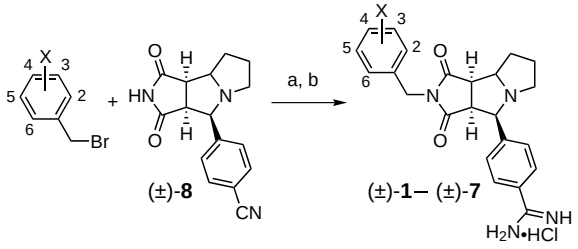


Figure 1. Schematic representation of the binding mode of the tricyclic inhibitors in the active site of thrombin. Only the (3a*S*,4*R* 8a*S*,8b*R*)-configured enantiomer is bound, as determined from the X-ray crystallographic analysis.[10] In addition to the catalytic triad flanked by the oxyanion hole, the active site can be described in terms of selectivity pocket (S1), a small proximal pocket (P), and a large hydrophobic distal (D) pocket.

Benzyl derivative (±)-**1**[10c] and a series of six mono-, di- and pentafluorinated inhibitors (±)-**2**–(±)-**7** (Scheme 1) were prepared by alkylation of (±)-**8** (which was obtained by 1,3-dipolar cycloaddition from maleimide, 4-formylbenzonitrile, and L-proline) followed by a Pinner reaction.[10c,13] Their activity against thrombin and trypsin was determined as previously described.[14] The results were surprising (Table 1): whereas five fluorinated inhibitors ($K_i = 0.27$ to 0.61 μM) showed a potency similar to that of non-fluorinated (±)-**1** ($K_i = 0.31$ μM), the *para*-monofluorinated derivative (±)-**4** was much more active, with $K_i = 0.057$ μM



Scheme 1. Synthesis of inhibitors (±)-**1** to (±)-**7**. a) K_2CO_3, [18]crown-6 (10%), toluene, reflux, 58–86%; b) 1. AcCl, MeOH, CH_2Cl_2, 5°C; 2. NH_3 in MeOH, 65°C, 40–79%.

[*] Dr. D. W. Banner, Prof. Dr. K. Müller, Dr. U. Obst Sander, A. D'Arcy, M. Stihle
Pharma Research Basel, Discovery Chemistry
F. Hoffmann–La Roche Ltd
4070 Basel (Switzerland)
Fax: (+41) 61-688-7408
E-mail: david.banner@roche.com
klaus.mueller@roche.com
Prof. Dr. F. Diederich, Dr. J. A. Olsen, P. Seiler
Laboratorium für Organische Chemie
ETH-Hönggerberg, HCI
8093 Zürich (Switzerland)
Fax: (+41) 1-632-1109
E-mail: diederich@org.chem.ethz.ch

[**] We thank Dr. T. Tschopp for the measurement of enzyme inhibition constants. J.A.O. thanks the Carlsberg Foundation for a fellowship. The research was supported by the ETH Research Council and F. Hoffmann–La Roche Ltd.

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

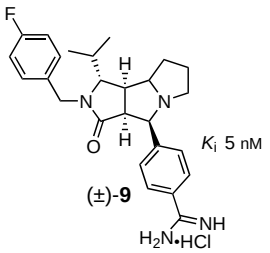
 DOI: 10.1002/ange.200351268

Table 1: Activities of the fluorinated thrombin inhibitors and selectivities with respect to trypsin.

Inhibitor	X (R^1)	K_i [μM][a]	Selectivity[b]
(±)-**1**	—	0.31	15
(±)-**2**	2-F	0.50	9.8
(±)-**3**	3-F	0.36	26
(±)-**4**	4-F	0.057	67
(±)-**5**	2,6-F_2	0.61	9.0
(±)-**6**	3,5-F_2	0.59	25
(±)-**7**	1,2,3,4,5-F_5	0.27	44

[a] The uncertainty of the measured K_i values is ±20%. [b] K_i(trypsin)/K_i(thrombin).

($\Delta\Delta G_{(\pm)\text{-}\mathbf{1}\rightarrow(\pm)\text{-}\mathbf{4}} = -1.05 \pm 0.17$ kcal mol^{-1}). Furthermore, this compound showed by far the best selectivity against trypsin in the series tested. We subsequently synthesized inhibitor (±)-**9**, with an isopropyl group to fill the P pocket, by a previously reported route.[10a–b] Compound (±)-**9** with a K_i value of 5 nM is to date our most potent thrombin inhibitor with a 413-fold selectivity over trypsin.



The binding mode of (±)-**4** at the thrombin active site was investigated by X-ray crystallography at 1.67-Å resolution (Figure 2).[15] Not surprisingly, only the 3a*S*,4*R*,8a*S*,8b*R* enantiomer is bound, as we had already shown for related thrombin inhibitors.[10a,b] Superimposition with a previously reported X-ray crystal structure of thrombin complexed with an inhibitor having a piperonyl group pointing into the D pocket[10a] revealed an identical binding mode for both inhibitors (see the Supporting Information). This result supports our modeling-based assumption that all fluorinated inhibitors have identical binding geometries.

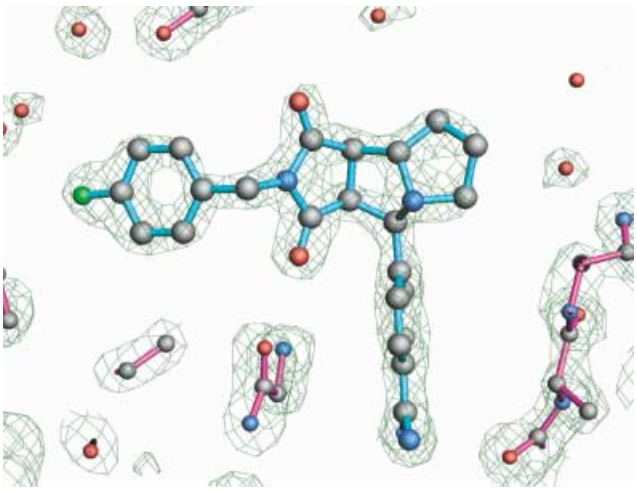


Figure 2. $2F_oF_c$ electron density contoured at 0.6 e Å^{-3}. All the inhibitor atoms are well-ordered. The slightly weaker electron density for the pyrrolidine ring hints at pseudorotation, although the temperature factors after constrained refinement have acceptably low values of approximately 20 Å^2.[15]

The F atom of (±)-**4** is in remarkably close contact with the H–C_α–C=O moiety of Asn 98 (Figure 3), as indicated by the short distances d(F···Cα) of 3.1 Å (d(F···H_α): 2.1 Å,

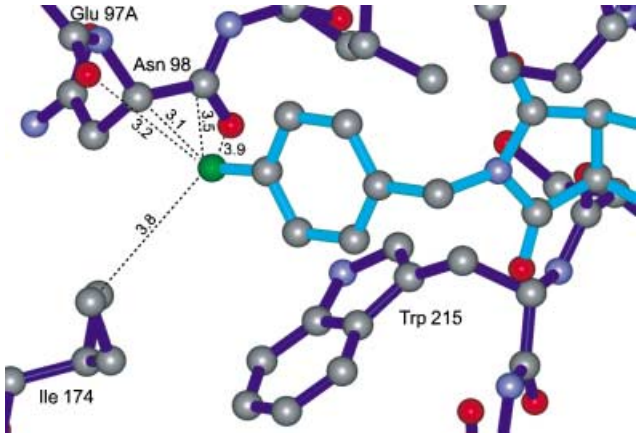
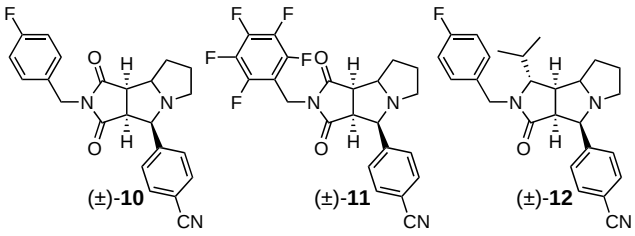


Figure 3. Binding mode of the 4-fluorobenzyl moiety in the D pocket of thrombin.[15] All F contacts with the protein within a distance of less than 4.0 Å are shown with dotted lines.

α(F···H–C_α): 157°) and d(F···C=O): 3.5 Å (α(F···C=O): 96°, α(F···C(0)–pseudotrigonal axis of the carbonyl system): 28°).[16] The ability of the H–C_α bond to act as an H-bond donor is increasingly being recognized, and calculations show that weak C=O···H–C_α H bonds contribute significantly in stabilizing the folded structure of proteins.[17] Desiraju and coworkers obtained evidence for weak C–F···H–C interactions from crystal structures of fluorobenzenes.[18] In addition, it was recently reported that C–F···H–C interactions stabilize duplex formation with RNA having unnatural fluorinated bases.[19] Whether the observed C–F···H–C_α interaction should be classified as a weak H bond is a matter of discussion.[20] The enhanced positive polarization of the H atom *ortho* to the F substituent presumably enhances the edge-to-face interaction of the inhibitor with Trp 215 (Figure 3).[21] We believe, however, that a particularly comfortable positioning of the F atom, with its direct polar interactions with the H–C_α–C=O moiety of Asn 98, is the major determinant for the large increase in activity seen for (±)-**4** relative to the other inhibitors in the series.

Organofluorine interactions with carbonyl groups, as seen in the thrombin complex with (±)-**4**, have to the best of our knowledge not been reported in the literature. We therefore analyzed the small-molecule X-ray crystal structures of the nitrile precursors (±)-**10**, (±)-**11**, and (±)-**12** to inhibitors (±)-**4**, (±)-**7**, and (±)-**9**, respectively, to obtain further insight into fluorophilic environments.[22]



A C–F···H–C_α–C(O) contact similar to the one in the thrombin complex was found in the crystal-packing environment of (±)-**12**, as indicted by d(F···H_α): 2.41 Å,[23] d(F···C_α): 3.37 Å, and α(F···H-C_α): 148° (Figure 4). The F atom is rather

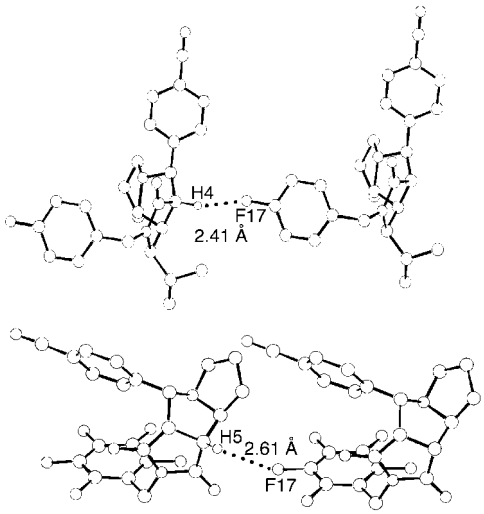


Figure 4. Short intermolecular C–F···H–C_α–C=O contacts[23] observed in the X-ray crystal packing of (±)-**11** (bottom) and (±)-**12** (top).[22]

distant from the C=O unit (d(F···C=O): 4.53 Å and α(F···C=O): 110°). Two similar C–F···H–C_α interactions were identified, though at longer F···H distances (2.65 and 2.74 Å), in the crystal of (±)-**10** (see the Supporting Information).[23] In the first case, the F atom is also in contact with a C=O group (d(F···C=O): 3.40 Å, d(F···O=C): 3.06 Å, α(F···C=O): 63°).

Analysis of the crystal packing of compound (±)-**11** revealed an additional C–F···H–C_α–C=O contact (d(F···H_α): 2.61 Å,[23] d(F···C_α): 3.27 Å, α(F···H-C_α): 119°, Figure 4). In this case the F atom is also in contact with the carbonyl group (d(F···C=O): 3.63 Å, d(F···O=C): 3.30 Å, α(F···C=O): 65°).

These findings prompted us to search the Cambridge Structural Database (CSD)[24] for intermolecular interactions between C–F and C=O moieties in crystals of fluorine-containing organic carbonyl derivatives in which the C=O group is flanked by one or two C, N, or O atoms. The search was limited to well-resolved (R factor < 7.5%, no disorder) neutral organic molecules, excluding metallorganic or polymeric compounds, with an intermolecular (nonbonded) distance between a C-bound F atom and a carbonyl C atom in the range of 2.5 Å < d_1(F···C=O) < 4.0 Å. The search resulted in 573 hits, with a total of 1091 occurrencies of close F···C=O contacts. There are quite a number of cases (43 occurrencies, Figure 5) with contact distances significantly below the sum of the van der Waals radii of C and F atoms (ca. 3.3 Å), which indicates there is an attractive interaction between the two polar groups.

An analysis of the relative orientation of the two interacting units in terms of the nonbonded distance d_1(F···C=O), the two angles α_1 and α_2 (Figure 5), as well as the pseudotorsional angle θ (C–F···C=O, Figure 5 and see the Supporting Information) reveals that the F atoms prefer a position close to the pseudotrigonal axis of the carbonyl unit

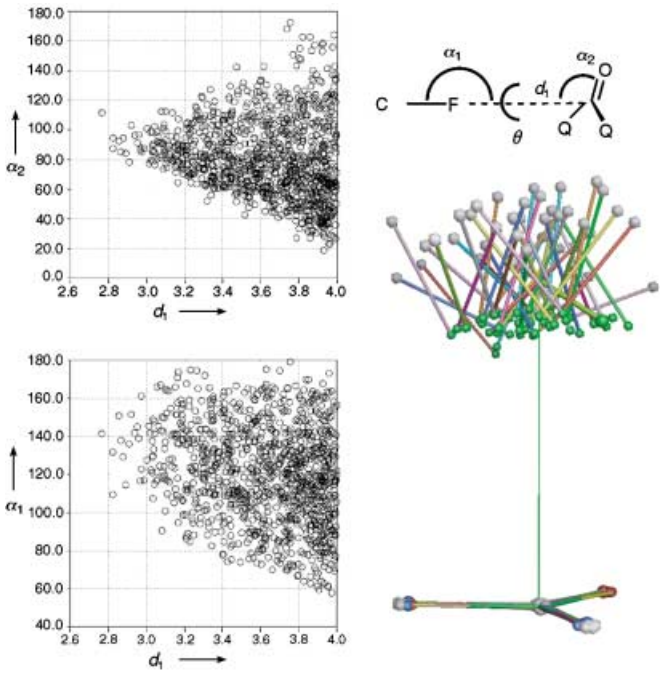


Figure 5. Scatterplots of α_1 versus d_1 (upper left) and α_2 versus d_1 (lower left) of 1091 occurrencies of close C–F···C=O contacts with 2.5 Å < d_1 < 4.0 Å obtained from 573 hits of the CSD search of fluorine-containing organic carbonyl derivatives (Q = C, N, O). Right: Superposition of 43 intermolecular C–F···Q(CO)Q subunits with 2.77 Å < d_1 < 3.09 Å extracted from the crystal structures and superimposed on the carbonyl unit (pointing backwards to the right). The thin vertical line marks the pseudotrigonal axis of the carbonyl system (norm of the Q(CO)Q plane).

at short contact distances. The C–F bond is generally inclined to the F–C axis with angles (α_1) adopting values typically between 100°–160°, but rarely 180°, at short to middle-range contact distances. There appears to be no preference in terms of *anti*, *syn*, or other orientations of the C–F unit relative to the C=O bond, even for the closest contacts. No systematic change of either the C–F or C=O bond lengths in response to the nonbonded distance d_1(F···C=O) could be found (see the Supporting Information). Thus, there are no indications of a covalent component in these C–F···C=O interactions, which are probably best described in terms of multipolar interactions between the intrinsically polar C–F and C=O units.

The crystal structures with short intermolecular contact distances of 2.9 Å < d_1(F···C=O) < 3.1 Å were examined in more detail using the modeling system Moloc:[16] the close

intermolecular contacts were identified, the relevant substructural units extracted from the crystal structures, and superimposed with respect to the carbonyl system. The plot (Figure 5, right) shows the narrow cone around the pseudo-trigonal axis of the carbonyl group (cone angle ca. 12°) within which the fluorine atom of a closely interacting C–F bond unit prefers to be located.[25]

In view of our observation that the fluorine atom in the thrombin inhibitor complex also appears to be in close contact to a C_α–H bond α next to the C=O group, we analyzed a subset of structures containing at least one such bond in terms of close contacts of the C–F unit to both the C=O and the C_α–H unit.

Interestingly, while there is an expected correlation between the nonbonded distances d_3(F···C_α) and d_1(F···C=O) (Figure 6, left), the scatter plot of d_2(F···H–C_α) versus d_1(F···C=O) (Figure 6, right) clearly shows that a close contact of the fluorine atom with the carbonyl group does not

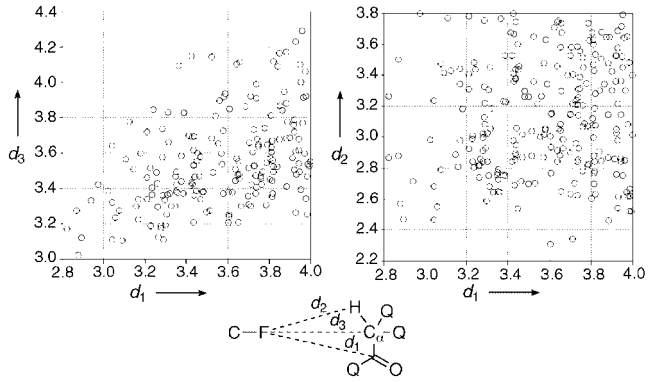


Figure 6. Scatterplots of nonbonded contact distances d_3(F···H–Cα) versus d_1 (left) and d_2(F···C_α) versus d_1 (right) for a total of 246 occurrencies in the subset of 132 crystal structures in which the C=O group is attached to at least one C atom carrying at least one H atom.

necessarily imply a close C–F···H–C_α contact. Conversely, there are clear cases where the F atom interacts predominantly with the C_α–H unit, but is somewhat more remote from the carbonyl group. Thus, there appears to be a range of possibilities for a C–F unit to engage in potentially attractive contacts. It is gratifying to see that the pair of values of 3.1 Å and 3.5 Å observed for the F···C_α and F···C=O nonbonded contact distances, respectively, in the thrombin inhibitor complex lies essentially in the domain of the scatter plot obtained for corresponding contacts in small-molecule crystal structures (Figure 6, left).

In conclusion, the combination of X-ray crystal-structure analysis of a protein–ligand complex, small-molecule X-ray crystallography, and database mining has for the first time shown that H–C_α–C=O fragments provide a pronounced fluorophilic environment. The high propensity of H–C_α–C=O units in the active sites of proteins suggests that such F interactions could be effectively exploited for enhancing ligand affinity or selectivity in structure-based design.

Received: February 24, 2003 [Z51268]

Keywords: fluorine · inhibitors · medicinal chemistry · molecular recognition · structure elucidation · thrombin

[1] 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.

[2] M. Schlosser, D. Michel, *Tetrahedron* **1996**, *52*, 99–108.

[3] J. Hernández-Trujillo, A. Vela, *J. Phys. Chem.* **1996**, *100*, 6524–6530.

[4] 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.

[5] J. A. Ippolito, D. W. Christianson, *Int. J. Biol. Macromol.* **1992**, *14*, 193–197.

[6] 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.

[7] a) C.-Y Kim, J. S. Chang, J. B. Doyon, T. A. 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.

[8] A. Kashima, Y. Inoue, S. Sugio, I. Maeda, T. Nose, Y. Shimohigashi, *Eur. J. Biochem.* **1998**, *255*, 12–23.

[9] F. T. T. Huque, K. Jones, R. A. Saunders, J. A. Platts, *J. Fluorine Chem.* **2002**, *115*, 119–128.

[10] 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.

[11] For recent examples of other thrombin inhibitors, see a) J. Z. Ho, T. S. Gibson, J. E. Semple, *Bioorg. Med. Chem. Lett.* **2002**, *12*, 743–748; b) S. Hanessian, E. Therrien, K. Granberg, I. Nilsson, *Bioorg. Med. Chem. Lett.* **2002**, *12*, 2907–2911; c) for a review, see J. B. M. Rewinkel, A. E. P. Adang, *Curr. Pharm. Des.* **1999**, *5*, 1043–1075.

[12] The term fluorine scan has been coined in a recent report on a computational fluorine scan of biotin, see B. Kuhn, P. A. Kollman, *J. Am. Chem. Soc.* **2000**, *122*, 3909–3916.

[13] All new compounds were fully characterized by m.p., IR, ^{1}H and ^{13}C NMR spectroscopy, MS, and HR-MS or elemental analysis.

[14] 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.

[15] Crystals of human α-thrombin of C_2 symmetry were grown in the presence of a carboxy-terminal hirudin peptide (des-amino Asp55 hirudin55–65) as the inhibitor in the anion-binding exosite, but with no active-site inhibitor. Compound (±)-**4** was diffused into the crystals for 6 h from a 2.5 mM solution containing 30% PEG3350, 100 mM NaCl, and 100 mM Na^+/K^+ pH 7.4 phosphate buffer. One crystal was flash-frozen in liquid N_2, and diffraction data were measured on a Delft Instruments FR591 X-ray generator, run at 50 kV/100 mA, equipped with an Osmic focussing monochromator and an Oxford Cryostream cooler run at 100 K. The Marresearch345 image plate detector was placed 120 mm from the crystal and scanned with 0.15 mm pixel size. Exposure times were 180 s for 1.0° frames, and a profuse amount of data was obtained by collecting 720 frames. Data were processed to 1.67-Å resolution using XDS (W. Kabsch, *J. Appl. Crystallogr.* **1993**, *26*, 795–800). Unit-cell

dimensions were $a = 70.69$, $b = 71.48$, $c = 72.80$ Å, and $\beta = 100.53°$. For 284109 observations of 38758 reflections, the merging R factor on intensities was 3.3% and in the outermost data shell (1.67 Å–1.74 Å) 14.9%, with $I/\sigma I = 33.5$ overall and 9.7 in the outermost shell. Model building with Moloc[16] and refinement with CNX (A. T. Brünger, P. D. Adams, G. M. Clore, W. L. DeLano, P. Gros, R. W. Grosse-Kunstleve, J.-S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson, G. L. Warren, *Acta Crystallogr. D* **1998**, *54*, 905–921) gave final overall crystallographic R factors of 17.9% (working) and 20.5% (free), with values in the outer shell of 22.8 and 24.9%, respectively, for 2296 protein atoms, 30 inhibitor atoms, one Na^+ ion, one Ca^{2+} ion, and 404 water molecules. The inhibitor density is very clear (Figure 2): the refined temperature factors for the inhibitor range from 14 (amidinium moiety) to 26 Å^2. PDB code: 10YT.

[16] a) P. R. Gerber, K. Müller, *J. Comput.-Aided Mater. Des.* **1995**, *9*, 251–268; b) Gerber Molecular Design (http://www.moloc.ch). The distances and angles were obtained with the Moloc modeling/visualization program. The position of the H_α atom of Asn98 is based on stereochemical consideration with a C···H distance of 1.08 Å.

[17] R. Vargas, J. Garza, D. A. Dixon, B. P. Hay, *J. Am. Chem. Soc.* **2000**, *122*, 4750–4755.

[18] V. R. Thalladi, H.-C. Weiss, D. Bläser, R. Boese, A. Nangia, G. R. Desiraju, *J. Am. Chem. Soc.* **1998**, *120*, 8702–8710.

[19] J. Parsch, J. W. Engels, *J. Am. Chem. Soc.* **2002**, *124*, 5664–5672.

[20] J. D. Dunitz, R. Taylor, *Chem. Eur. J.* **1997**, *3*, 89–98.

[21] A. Matsushima, T. Fujita, T. Nose, Y. Shimohigashi, *J. Biochem.* **2000**, *128*, 225–232.

[22] CCDC 203566–203568 ((±)-**10**–(±)-**12**) 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) 1223-336-033; or deposit@ccdc.cam.ac.uk). Further details are found as Supporting Information.

[23] The H positions are based on stereochemical consideration with a C···H distance of 1.08 Å.

[24] CSD Version 5.24, November **2002**, which holds over 270000 small-molecule crystal structures.

[25] The F···C=O interaction bears some structural resemblance to the interaction of nucleophiles with the carbonyl groups investigated: B. Bürgi, J. D. Dunitz, *Acc. Chem. Res.* **1983**, *16*, 153–161.

