Protease Inhibitors: Synthesis and QSAR Study of Novel Classes of Nonbasic Thrombin Inhibitors Incorporating Sulfonylguanidine and *O*-Methylsulfonylisourea Moieties at P1

Claudiu T. Supuran,*,† Andrea Scozzafava,† Fabrizio Briganti,† and Brian W. Clare*,‡

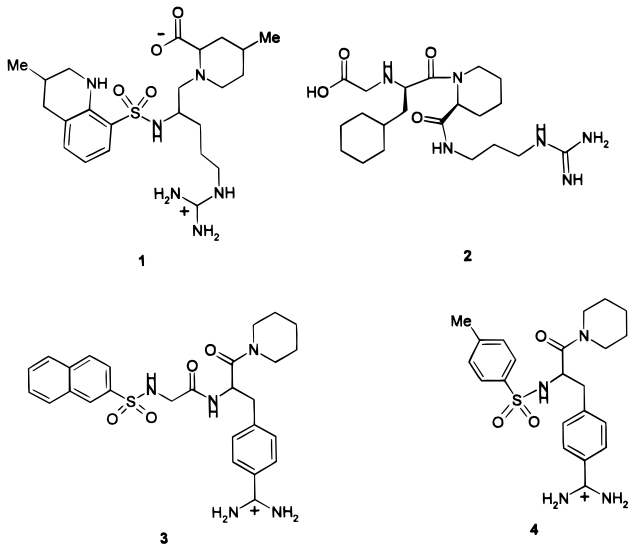
Laboratorio di Chimica Inorganica e Bioinorganica, Università degli Studi, Via Gino Capponi 7, I-50121 Florence, Italy, and Department of Chemistry, The University of Western Australia, Nedlands, Western Australia, Australia 6907

Received July 16, 1999

Using benzamidine as a lead molecule, two series of alkyl/aralkyl/arylsulfonylguanidines/sulfonyl-*O*-methylisoureas have been prepared and assayed as inhibitors of two serine proteases, thrombin and trypsin. The study showed that sulfaguanidine and its corresponding *O*-methylisourea derivative possess moderate but intrinsically selective thrombin inhibitory properties, with K_I's around 100 nM against thrombin and 1350−1500 nM against trypsin. Further elaboration of these two molecules afforded compounds that inhibited thrombin with K_I's in the range of 12−50 nM, whereas affinity for trypsin remained relatively low. Such compounds were obtained by attaching benzyloxycarbonyl- or 4-toluenesulfonylureido-protected amino acids (such as L- and D-Phe or L-Pro) or dipeptides (such as Phe-Pro, Gly-His, β-Ala-His, or Pro-Gly) to the two leads mentioned above, sulfaguanidine and 4-aminobenzenesulfonyl-*O*-methylisourea. Thus, the present study proposes two novel approaches for the preparation of high-affinity, specific thrombin inhibitors: two novel S1 anchoring moieties in the already large family of arginine/amidine-based inhibitors and novel peptidomimetic scaffolds obtained by incorporating tosylureido amino acids in the hydrophobic binding site(s). The first one is important for obtaining bioavailable thrombin inhibitors, devoid of the high basicity of the commonly used arginine/amidine-based inhibitors, whereas the second one may lead to improved water solubility of such compounds due to facilitated metal (sodium) salts formation (at the relatively acidic SO_2NHCO protons) as well as increased stability at hydrolysis (in vivo). A QSAR study also explained the activity in terms of global properties of the molecules, electronic properties of the sulfonylguanidine/sulfonylisourea moiety, and novel descriptors, the frontier orbital phase angles (FOPA), that account for the directions of the nodes in the π orbitals in the aromatic portion of those of the drugs in which the sulfonyl group was bound to a benzene ring. For thrombin inhibition, the size of the molecule was the dominant influence, while for trypsin inhibition the FOPA was the principal determinant of activity. The dependence of activity on the FOPA variables is perhaps the clearest example of a quantum effect in pharmacology and suggests a promising new tool for drug design.

Introduction

Thrombin (EC 3.4.21.5) has become an important target for drug design in recent years, in the search for low-molecular-weight, potent, and selective inhibitors with applications as diagnostic and therapeutic agents for the increasingly common thrombotic diseases.[1−6] Although a large number of potent active site-directed thrombin inhibitors such as peptide aldehydes,[7,8] boronates,[9] and benzamidine-[2,10,11] or arginine/guanidine-derived[12] inhibitors were reported, none meet all the criteria needed for an ideal antithrombotic drug.[2,13] Thus, the largest majority of the presently available low-molecular-weight inhibitors, such as argatroban (MQPA) **(1)**,[1] inogatran **(2)**,[6] NAPAP **(3)**,[15] 4-TAPAP **(4)**, or its 3-amidino isomer 3-TAPAP **(5)**,[2,15] are poorly bioavailable, due to their high basicity, connected with the presence of guanidino/amidino moieties in their molecule, are not absorbable orally, or are rapidly eliminated from the circulation, mainly due to their peptidic nature. Although recently some nonbasic S1 anchoring groups have been incorporated in the molecules of some



* Correspondence authors. For C.T.S.: fax, +39-055-2757555; e-mail, cts@bio.chim.unifi.it. For B.W.C.: phone, +61-8-9380 3140; fax, +61-8-9380 1005; e-mail, bwc@crystal.uwa.edu.au.
† Università degli Studi.
‡ The University of Western Australia.

10.1021/jm9903693 CCC: $19.00
Published on Web 04/18/2000

thrombin inhibitors,[3,5,16] the presence of guanidino/benzamidino moieties in such compounds is critical, since it is by means of the interaction of these highly polar groups with Asp 189, the central amino acid residue from the specificity pocket, that the enzyme−inhibitor adduct is initially formed (obviously, a lot of other secondary interactions are responsible for the formation of high-affinity adducts between thrombin and its inhibitors).[3−5,10−12] To exploit the intrinsically high affinity of guanidino/benzamidino-containing inhibitors for the thrombin active site, but also to avoid undesired properties connected with their too high basicity, we propose here a novel approach for designing tight-binding inhibitors, by using sulfonylguanidino (and some of its variants) moieties as anchoring groups to the specificity S1 pocket. Obviously, the presence of the SO_2 group in the neighborhood of the guanidino moiety strongly reduces the basicity of the latter, presumably without precluding the binding of inhibitors within the enzyme active site.

In this paper we report the preparation and serine protease inhibitory properties (against human thrombin and human trypsin) of two series of alkyl/aralkyl- or arylsulfonylguanidines/sulfonyl-*O*-methylisoureas. The most promising compounds (from the point of view of their thrombin inhibitory properties, as well as that of their specificity for thrombin over trypsin) were then further elaborated, by attaching amino acyl/dipeptidyl moieties into their molecule, leading thus to high-affinity inhibitors, with potencies of the same order of magnitude as those of the clinically used compounds argatroban (MQPA) **(1)**[14] and inogatran **(2)**.[6]

Furthermore, a QSAR study of the new series of obtained thrombin inhibitors should throw light on the molecular features required for a good inhibitor and also on the molecular environment of the drug in the binding site and the mechanism of action. To this end we have carried out AM1 molecular orbital (MO) calculations on the drugs and regression analyses to relate the measured activity of the drugs to the calculated MO indices. The chosen indices include charges and superdelocalizabilities of atoms of the sulfonyl and guanidine/isourea moieties, dipole moments, orbital energies, polarizabilities, lipophilicities, aqueous solvation energies, and frontier orbital phase angles (FOPA).

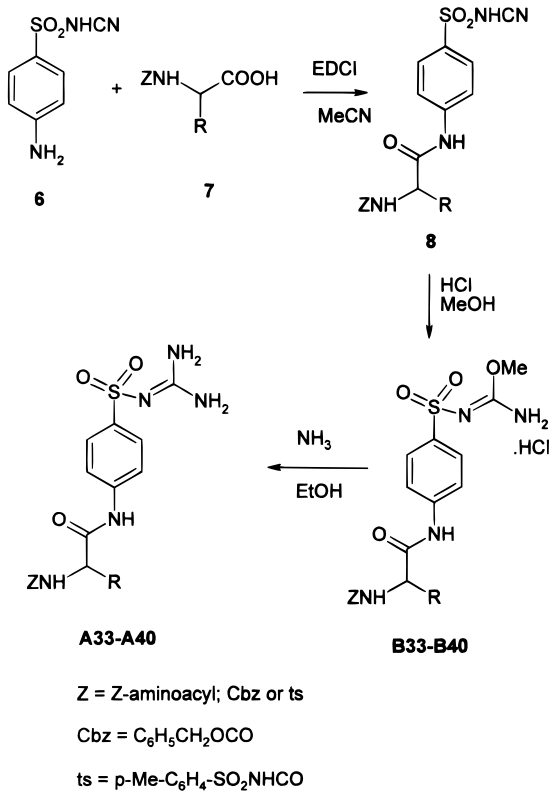
Results

Synthesis. Compounds prepared by reaction of alkyl-, aralkyl-, or arylsulfonyl halides with guanidine or *O*-methylisourea, of types **A,B(1−40)**, are shown in Tables 1−4.

Nonexceptional routine synthetic procedures have been used for the reactions of sulfonyl halides with nucleophiles (for the preparation of compounds **A,B(1−32)**),[17,18] whereas for attaching the amino acyl moieties to the leads **A14** and **B14**, the procedure is outlined in Scheme 1.

The key intermediate, sulfanilylcyanamide **6**,[19,20] was coupled with *N*-Cbz- or *N*-tosylureido-protected amino acids/dipeptides[21] **7** in the presence of EDCI or diisopropylcarbodiimide as condensing agents.[22] The intermediate sulfanilylcyanamide derivatives **8** were treated with anhydrous HCl in methanol, to afford the sulfonyl-*O*-methylisoureas **B33−B40**, which by reaction with alcoholic ammonia led to the sulfonylguanidines **A33−A40**.[2b,19]

Scheme 1



Serine Protease Inhibitory Activity. Inhibition data against two serine proteases, human thrombin and human trypsin, are shown in Tables 1−4. The chromogenic substrate Chromozym TH (Ts-Gly-Pro-Arg-*p*-nitroanilide) was used in the assay, with the spectrophotometric method of Lottenberg et al.[23] Inhibition data with the standard inhibitors **1−3** are also provided for comparison in Tables 1 and 2.

QSAR Calculations. In addition to the compounds of Tables 1−4, we include the sulfonylaminoguanidines which we have described previously.[24] The formulas and activity of these compounds are given in Tables 5 and 6. Only those molecules with a benzene ring bound to the sulfonamido sulfur were considered.

The structures were set up using the program Hyperchem[25] and initially optimized using the MM+ molecular mechanics option, varying in particular the dihedral angles of the guanidine/aminoguanidine/isourea moiety to obtain a global minimum. The peptide parts of molecules **33−40** were added assuming α-helical stereochemistry, and the completed molecules were then optimized with MOPAC 6,[26] using the AM1 Hamiltonian,[27] the EF (eigenvector following) optimization option, and the PRECISE option to ensure complete convergence. The polarizability was also computed during this step. The molecules were oriented such that the center of the benzene ring was at the origin, atom 1 (to which the sulfonamido S was bonded) was on the positive *X* axis, and the midpoint of atoms 2 and 3 was in the *X*−*Y* plane. The *Z* coordinate of the guanidine/isourea carbon was positive. A single-point calculation was then run on this geometry using MOPAC 93,[28] employing the COSMO[29] model to approximate the

Table 1. Sulfonylguanidines **A1**−**A32** Prepared in the Present Study, with Their Inhibition Data Against Human Thrombin and Human Trypsin

$RSO_2N{=}C(NH_2)_2$

A1-A32

compd **A**	R	K_I (nM)[a] thrombin	trypsin
1	Me_2N	1200	2890
2	$PhCH_2$	330	1350
3	CF_3	890	1275
4	*p*-F-C_6H_4	240	1090
5	*p*-Cl-C_6H_4	225	1170
6	*p*-Br-C_6H_4	220	1230
7	*p*-I-C_6H_4	210	1350
8	*p*-CH_3-C_6H_4	290	1810
9	*p*-O_2N-C_6H_4	180	975
10	*m*-O_2N-C_6H_4	190	1100
11	*o*-O_2N-C_6H_4	320	1850
12	3-Cl-4-O_2N-C_6H_3	160	990
13	*p*-AcNH-C_6H_4	195	1070
14	*p*-H_2N-C_6H_4	95	1350
15	*m*-H_2N-C_6H_4	107	1145
16	C_6F_5	146	1350
17	*o*-HOOC-C_6H_4	240	1445
18	*m*-HOOC-C_6H_4	121	1500
19	*p*-HOOC-C_6H_4	104	1235
20	*o*-HOOC-C_6Br_4	225	1250
21	*p*-CH_3O-C_6H_4	240	1320
22	2,4,6-$(CH_3)_3$-C_6H_2	255	1450
23	4-CH_3O-3-H_2N-C_6H_3	103	1080
24	2-HO-3,5-Cl_2-C_6H_2	152	1420
25	4-Me_2N-C_6H_4-N=N-C_6H_4	134	1245
26	5-dimethylamino-1-naphthyl	120	1150
27	1-naphthyl	136	1230
28	2-naphthyl	132	1300
29	*n*-C_4F_9	520	2100
30	*n*-C_8F_{17}	360	1950
31	2-thienyl	185	1320
32	camphor-10-yl	320	980
	benzamidine	300	450

[a] K_I values were obtained from Dixon plots using a linear regression program, from at least three different assays. Errors (data not shown) were ±5−10 % of the shown values.

properties of the molecule in solution in water, using a dielectric constant of 72.7. The FOPA Θ_H and Θ_L were calculated for the aromatic molecules,[30] and the Mulliken charges Q and the Fukui electrophilic and nucleophilic superdelocalizabilities S^E and S^N [31,32] were calculated for each atom in the molecules. For those compounds with a carboxyl group (**A**−**C, 17**−**20**) a preliminary Pallas calculation indicated that the compound was ionized at physiological pH, and the compound was calculated as an anion. All others were calculated as neutral molecules. The octanol−water partition constants were calculated using the program ClogP.[33]

The descriptors chosen for examination in the QSAR study included, for the sulfonamido moiety, Q, S^E, and S^N for the sulfonamide S, O, and N and the guanidine C for all of the molecules, the mean values of these for the guanidine and aminoguanidine NH_2 N and NH_2 H, the value for the isourea NH_2 N, and the mean value for the isourea NH_2 H. The symbol Na will refer to the guanidine nitrogen and Nb to the sulfonamide nitrogen. Thus, for example, Q_{Na} refers to the Mulliken charge of the guanidine/aminoguanidine/isourea NH_2 nitrogen atom, or the mean if there is more than one. In addition, E_{SH}, E_H, E_L, and E_{SL} refer to the second highest, highest, lowest, and second lowest π orbital of the benzene ring. Such a π orbital was recognized in the output by 1. The sum of squares of all of the p_z atomic orbital coefficients

Table 2. Derivatives **A33**−**A40** Obtained from Sulfaguanidine **A14** as Lead, with Their Inhibition Data Against Human Thrombin and Human Trypsin (standards **1**−**3** also included)

p-XNH-C_6H_4-$SO_2N{=}C(NH_2)_2$

A33-A40

compd	X[c]	K_I (nM)[a] thrombin	trypsin
A33	Cbz-D-Phe	54	1285
A34	ts-D-Phe	43	1250
A35	ts-L-Pro	48	1445
A36	ts-D-PhePro	12	1315
A37	Cbz-D-PhePro	13	1360
A38	ts-GlyHis	18	1455
A39	ts-*β*-AlaHis	15	1350
A40	ts-L-ProGly	21	1400
1	argatroban[b]	19	
2	inogatran	15	540
3	NAPAP	6.5	690

[a] K_I values were obtained from Dixon plots using a linear regression program, from at least three different assays. Errors (data not shown) were ±5−10 % of the shown values. [b] From ref 5a. [c] Cbz = $PhCH_2OCO$, ts = *p*-$MeC_6H_4SO_2NHCO$; these groups acylate the amino-terminal H_2N moiety. When configuration is not specified, L-amino acid moieties where employed. The usual polypeptide formalism is used: the amino-terminal residue is written first (and is always protected by either the Cbz or the ts moieties), whereas the carboxy-terminal residue is acylating the sulfaguanidine N-4 amino group.

Table 3. Sulfonyl-*O*-methylisoureas **B1**−**B32** Prepared in the Present Study, with Their Inhibition Data Against Human Thrombin and Human Trypsin

$RSO_2N{=}C(NH_2)OMe$

B1-B32

compd **B**	R	K_I (nM)[a] thrombin	trypsin
1	Me_2N	1365	3300
2	$PhCH_2$	450	1900
3	CF_3	895	1450
4	*p*-F-C_6H_4	280	1200
5	*p*-Cl-C_6H_4	266	1325
6	*p*-Br-C_6H_4	265	1370
7	*p*-I-C_6H_4	243	1415
8	*p*-CH_3-C_6H_4	328	2150
9	*p*-O_2N-C_6H_4	189	1235
10	*m*-O_2N-C_6H_4	213	1375
11	*o*-O_2N-C_6H_4	355	2340
12	3-Cl-4-O_2N-C_6H_3	177	1200
13	*p*-AcNH-C_6H_4	202	1345
14	*p*-H_2N-C_6H_4	106	1580
15	*m*-H_2N-C_6H_4	99	1370
16	C_6F_5	153	1425
17	*o*-HOOC-C_6H_4	325	1550
18	*m*-HOOC-C_6H_4	197	1595
19	*p*-HOOC-C_6H_4	133	1340
20	*o*-HOOC-C_6Br_4	239	1300
21	*p*-CH_3O-C_6H_4	276	1460
22	2,4,6-$(CH_3)_3$-C_6H_2	348	1785
23	4-CH_3O-3-H_2N-C_6H_3	96	1235
24	2-HO-3,5-Cl_2-C_6H_2	170	1550
25	4-Me_2N-C_6H_4-N=N-C_6H_4	169	1350
26	5-dimethylamino-1-naphthyl	138	1375
27	1-naphthyl	154	1425
28	2-naphthyl	147	1445
29	*n*-C_4F_9	565	2180
30	*n*-C_8F_{17}	403	2135
31	2-thienyl	213	1350
32	camphor-10-yl	336	1075

[a] K_I values were obtained from Dixon plots using a linear regression program, from at least three different assays. Errors (data not shown) were ±5−10 % of the shown values.

in the molecular orbital exceeded 0.5, 2. There were no large coefficients on any s, p_x, or p_y atomic orbitals, and

Table 4. Derivatives **B33**−**B40** Obtained from Sulfanilyl-*O*-methylisourea **B14** as Lead, with Their Inhibition Data Against Human Thrombin and Human Trypsin

p-XNH-C_6H_4-SO_2N=C(NH_2)OMe

B33-B40

compd **B**	X[b]	K_I (nM)[a] thrombin	trypsin
33	Cbz-D-Phe	62	1425
34	ts-D-Phe	60	1320
35	ts-L-Pro	63	1500
36	ts-D-PhePro	18	1420
37	Cbz-D-PhePro	16	1435
38	ts-GlyHis	21	1550
39	ts-β-AlaHis	19	1420
40	ts-L-ProGly	27	1540

[a] K_I values were obtained from Dixon plots using a linear regression program, from at least three different assays. Errors (data not shown) were ±5−10 % of the shown values. [b] Cbz = $PhCH_2OCO$, ts = *p*-$MeC_6H_4SO_2NHCO$; these groups acylate the amino-terminal H_2N moiety. When configuration is not specified, it means that L-amino acid moieties where employed. The usual polypeptide formalism is used: the amino-terminal residue is written first (and is always protected by either the Cbz or the ts moieties), whereas the carboxy-terminal residue is acylating the sulfanilyl-*O*-methylisourea N-4 amino group.

Table 5. Sulfonylaminoguanidines **C4**−**C28** Prepared in the Present Study, with Their Inhibition Data Against Human Thrombin and Human Trypsin

$RSO_2NHN=C(NH_2)_2$

C4-C28

compd **C**	R	K_I (nM)[a] thrombin	trypsin
4	*p*-F-C_6H_4	225	1025
5	*p*-Cl-C_6H_4	212	1100
6	*p*-Br-C_6H_4	203	1215
7	*p*-I-C_6H_4	177	1300
8	*p*-CH_3-C_6H_4	270	1775
9	*p*-O_2N-C_6H_4	166	990
10	*m*-O_2N-C_6H_4	170	1235
11	*o*-O_2N-C_6H_4	324	1800
12	3-Cl-4-O_2N-C_6H_3	154	1010
13	*p*-AcNH-C_6H_4	172	1025
14	*p*-H_2N-C_6H_4	91	1425
15	*m*-H_2N-C_6H_4	88	1400
16	C_6F_5	123	1350
17	*o*-HOOC-C_6H_4	205	1400
18	*m*-HOOC-C_6H_4	112	1520
19	*p*-HOOC-C_6H_4	97	1335
20	o-HOOC-C_6Br_4	213	1200
21	*p*-CH_3O-C_6H_4	227	1275
22	2,4,6-$(CH_3)_3$-C_6H_2	219	1345
23	4-CH_3O-3-H_2N-C_6H_3	219	1100
24	2-HO-3,5-Cl_2-C_6H_2	98	1100
25	4-Me_2N-C_6H_4-N=N-C_6H_4	139	1355
26	5-dimethylamino-1-naphthyl	130	1200
27	1-naphthyl	125	1200
28	2-naphthyl	129	1285

[a] K_I values were obtained from Dixon plots using a linear regression program, from at least three different assays. Errors (data not shown) were ±5−10 % of the shown values.

3. There were large coefficients on p_z atomic orbitals on the benzene ring bonded to the sulfonamide moiety. This isolated the particular molecular orbitals corresponding to the degenerate HOMO and LUMO of benzene approximating to the aromatic part of the drug molecule. The FOPA descriptors were calculated from these six coefficients as described previously,[30] numbering as 1 the atom bonded to the sulfonamide S. The splitting energies Δ_H and Δ_L were defined as $E_H - E_{SH}$ and $E_{SL} - E_L$, respectively. Two indicator variables I_A and I_B were defined such that $I_B = 1$ for isoureas and $I_B = 0$ otherwise and $I_A = 1$ for aminoguanidines and $I_A = 0$ otherwise. The lipophilicity log *P* could be calculated explicitly only for the guanidines. For the isoureas and aminoguanidines, it will differ from the result for the guanidines by a constant, and this constant will be incorporated into I_A and I_B. The polarizabilities Π_{xx}, Π_{yy}, and Π_{zz} are referred to the inertial axes and the dipole moment components μ_x, μ_y, and μ_z to the input orientation. The computed solvation energy ΔH_S is the difference between the heat of formation calculated by the COSMO method and the vacuum calculation. The van der Waals surface area A_w was obtained from the MOPAC 93 calculation and served as a measure of molecular size. The mean absolute Mulliken charge Q_m and the local dipole index D_l served as measures of charge separation in the molecule.

Table 6. Derivatives **C33**−**C40** Obtained from Sulfanilylaminoguanidine **C14** as Lead, with Their Inhibition Data Against Human Thrombin and Human Trypsin

p-XNH-C_6H_4-$SO_2NHN=C(NH_2)_2$

C33-C40

compd **C**	X[b]	K_I (nM)[a] thrombin	trypsin
33	Cbz-D-Phe	50	1315
34	ts-D-Phe	44	1340
35	ts-L-Pro	43	1530
36	ts-D-PhePro	10	1450
37	Cbz-D-PhePro	11	1400
38	ts-GlyHis	15	1555
39	ts-β-AlaHis	10	1390
40	ts-L-ProGly	15	1560

[a] K_I values were obtained from Dixon plots using a linear regression program, from at least three different assays. Errors (data not shown) were ±5−10 % of the shown values. [b] Cbz = $PhCH_2OCO$, ts = *p*-$MeC_6H_4SO_2NHCO$; these groups acylate the amino-terminal H_2N moiety. When configuration is not specified, L-amino acid moieties where employed. The usual polypeptide formalism is used: the amino-terminal residue is written first (and is always protected by either the Cbz or the ts moieties), whereas the carboxy-terminal residue is acylating the sulfaguanidine N-4 amino group.

Statistical calculations were carried out with the BMDP statistical package.[34] Regressions were done with the all possible subsets method,[35] with control of collinearity being exercised by running principal components analysis on the subsets of variables. The quantity[36]

$$\Lambda = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{\lambda_i}$$

was calculated, where N is the number of variables in the subset and λ_i are the eigenvalues of the correlation matrix of the subset of independent variables. A value of Λ less than 5 indicates no significant collinearity. If in any subset of interest Λ exceeded 5, the eigenvector matrix of the correlation matrix was examined. For the eigenvector corresponding to the smallest eigenvalue, large loadings indicated a collinearity between the corresponding variables. One or more such variables were excluded from consideration, or two or more variables combined, and the entire procedure was repeated until Λ values were obtained that indicated that collinearity was absent. Combined variables in-

cluded $\Pi = (\Pi_{xx} + \Pi_{yy} + \Pi_{zz})/3$, $E_{L-H} = E_L - E_H$, and $E_{L+H} = E_L + E_H$. Subsets of variables meeting these criteria are presented.

When selecting a subset of variables from a pool, the statistical significance tends to be inflated due to lack of consideration of the multitude of possible models, as shown by Topliss et al.[37,38] In the present study, this inflation of significance was controlled by the calculation of many stepwise regressions using the Effroymson algorithm[39] with the dependent variable randomly reassigned. Statistical significance was calculated from the distribution of R, as described previously.[40,41]

Discussion

Chemistry. The lead molecule for obtaining novel types of thrombin inhibitors considered by us was benzamidine **9**, one of the simplest such compounds, which possesses an inhibition constant $K_I = 300$ nM against human thrombin; moreover, the X-ray crystallographic structure for the complex of benzamidine with this enzyme has recently been reported (PDB entry: 1DWB).[42] From the X-ray data it was observed that the amidino moiety of the inhibitor is anchored to the S1 specificity pocket of the enzyme, interacting electrostatically and by means of hydrogen bonds with Asp 189. Several other van der Waals contacts between the inhibitor molecule and the enzyme were also evidenced.[42] Obviously, benzamidine is a weak thrombin inhibitor, since the binding energy is only gained due to the strong electrostatic interaction of the carboxylate of Asp 189 and the positively charged amidino moiety. On the other hand, as already mentioned in the Introduction, the amidino moiety possesses too high a basicity for allowing the formation of bioavailable enzyme inhibitors, and it appeared thus of great interest to elaborate nonbasic variants of this attractive thrombin anchoring group. The sulfonylguanidino moiety (as well as some of its variants, such as for instance the sulfonyl-*O*-methylisourea group) appeared as an attractive candidate for such a purpose, since the presence of the SO_2 moiety in the neighborhood of the strong base, guanidine, should drastically weaken its basicity. The second moiety on the other hand ($SO_2N{=}C(OMe)NH_2$) is much less basic than the guanidino group. Such modified anchoring groups should not presumably interfere with the binding of the inhibitor to the enzyme, since the hydrogen-bonding donor/acceptor properties as well as the possibility to interact electrostatically with the enzyme for the compounds incorporating them should not differ too much from those of the classical amidino/guanidino-based inhibitors of types **1**−**5** or **9**. Since sulfonylguanidines or sulfonyl-*O*-methylisoureas possess a large number of possible tautomeric forms, and this factor might be a critical one for the binding of such a compound to thrombin, we performed MOPAC as well as AM1 calculations in order to identify the most stable tautomers.

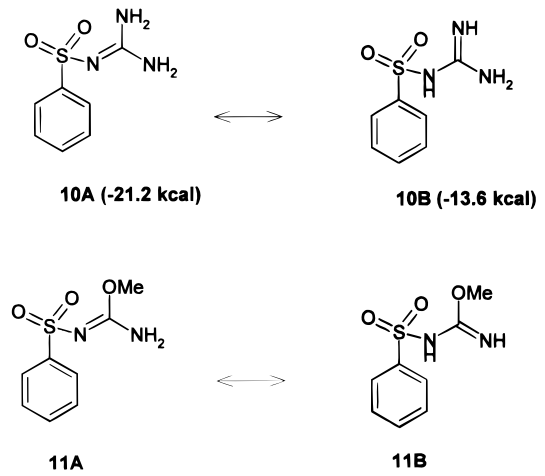
pK_a values for the amidino/guanidino as well as sulfonamido moieties of some of the newly synthesized serine protease inhibitors and standard compounds such as inogatran, argatroban, and NAPAP (Table 7) prove that the approach proposed here for reducing the basicity of such an enzyme inhibitor is a successful one. Thus, unlike the highly basic guanidines/amidines of types **1**−**3** (pK_a's around 12.3−12.6), sulfaguanidine **A14** and some of its derivatives reported here have pK_a values of the guanidino moiety around 7.9−8.4, being around 10^4 times less basic than the previously mentioned derivatives. Furthermore, due to the presence of the sulfonyl moiety in their molecules, these compounds also possess a weakly acidic character, with another ionization step around the pK_a value of 6.2−7.3, due to the loss of the SO_2NH proton. These features should positively influence the pharmacological profile of a thrombin inhibitor of the type described here.

Table 7. Experimentally Determined pK_a Values of Some Thrombin Inhibitors

inhibitor	pK_a[a] guanidino/amidino/isourea moiety	pK_a[a] SO_2NH moiety
1, argatroban[b]	12.5	
2, inogatran[b]	12.3	
3, NAPAP	12.6	
A14, sulfaguanidine	8.4	7.0
B14, sulfanilyl-*O*-Me-isourea	cannot be determined	6.2
C14, sulfanilylaminoguanidine	7.9	7.3

[a] pK_a's were determined in EtOH−water, 30% (v/v), as described in the Experimental section. [b] From ref 8.

Scheme 2



Calculations were done with Gaussian 94[43] and MOPAC 93.[26] For the molecules phenylsulfonylguanidine, phenylsulfonyl-O-methylpseudourea, and phenylsulfonylaminoguanidine, two tautomeric structures of each (shown in Scheme 2) were generated using Hyperchem 5.1[25] and optimized using Hyperchem's MM+ molecular mechanics optimizer. These geometries were further optimized with MOPAC 93 using the AM1 Hamiltonian.[27] The geometries thus obtained were used in one-point calculations to obtain the heats of formation in vacuo and in a medium of dielectric constant 72.7 using the COSMO[29] approximation (by MOPAC) and with the B3LYP method[44,45] and in the latter medium using SCIPCM[46] (by Gaussian). pK_a values were calculated using the Pallas software.[47]

The results are shown in Table 8 for the stabilities and Table 9 for the pK_a's. For both the density functional and AM1 methods, solvation increased the stability of structure 1 over structure 2. For the sulfonylguanidine and sulfonyl-*O*-methylpseudourea, both methods predicted greater stability for structure 1, but for the sulfonylaminoguanidine, AM1 predicted greater stability

Table 8. Stability Results[a]

compound	AM1/vacuum	AM1/COSMO	DF/vacuum	DF/SCIPCM
guanidine/tautomer 1	−21.44	−63.12	−984.962005	−984.985592
guanidine/tautomer 2	−13.75	−49.48	−984.952926	−984.969578
difference (1 − 2)	**−7.69**	**−13.64**	**−5.70**	**−10.05**
isourea/tautomer 1	−64.00	−101.32	−1044.117509	−1044.136993
isourea/tautomer 2	−56.36	−89.98	−1044.115495	−1044.131326
difference (1 − 2)	**−7.64**	**−11.34**	**−1.26**	**−3.56**
aminoguanidine/tautomer 1	0.47	−38.42	−1040.254819	−1040.278149
aminoguanidine/tautomer 2	10.95	−28.77	−1040.262145	−1040.282850
difference (1 − 2)	**−10.48**	**−9.65**	**+4.60**	**+2.95**

[a] DF = B3LYP/6-31G*//AM1. The AM1 results are in kcal and the density functional results in Hartrees, except for the differences (in bold) which are in kcal throughout.

Table 9. pK_a Results

compound	pK_a Pallas
guanidine/tautomer 1	*a*
guanidine/tautomer 2	2.3
isourea/tautomer 1	−6.82
isourea/tautomer 2	10.45, −4.92
aminoguanidine/tautomer 1	*a*
aminoguanidine/tautomer 2	6.84

[a] No acidic or basic groups found within aqeous range.

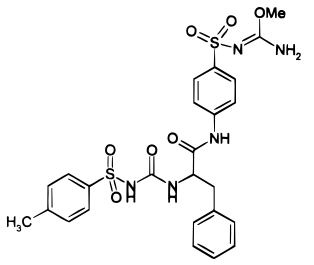
for structure 1 but the density functional method predicted the reverse, albeit with a fairly small difference.

Pallas predicts that in all cases, structure 1 will not protonate under aqueous conditions. Structure 2 for the sulfonylguanidine will protonate only at low pH, but for the sulfonylaminoguanidine it will undergo protonation close to physiological pH.

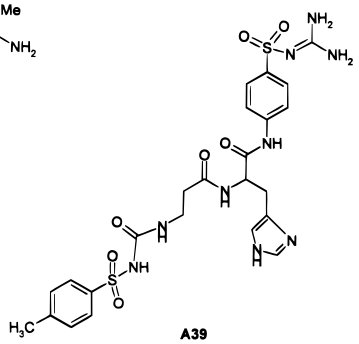
The tautomer of type **10A** is much more stable than the tautomer **10B** (Scheme 2) in the case of benzenesulfonylguanidine. The same situation is true for the corresponding benzenesulfonyl-*O*-methylisoureas **11A**,**B** (Scheme 2).

It is obvious from the above data that the symmetric nature of the favored tautomer **10A** should enable strong interactions with the carboxylate moiety of Asp 189 and the presumable formations of high-affinity E−I adducts. The same situation should apply for the formation of sulfonylisourea derivatives, although the pattern of hydrogen bonds/electrostatic interactions should be different in this case. In this context one should also mention that recently Salemme et al.[5b] have shown that the 4-hydroxy-3-methoxybenzyl group is capable of interacting with the guanidino-binding site of thrombin (when attached to a D-Phe-Pro scaffold). This last anchoring group is relatively similar with the sulfonyl-*O*-methylisourea moiety evidenced by us in this paper as an efficient S1 anchoring moiety. Thus, two series of sulfonylguanidines **A1**−**A32** and sulfonyl-*O*-methylisoureas **B1**−**B32** were prepared in order to test this hypothesis (Tables 1 and 3). These compounds were obtained by simple reactions of alkyl/aralkyl/aryl/hetarylsulfonyl halides or sulfonic acid anhydrides with nucleophiles such as guanidine, *O*-methylisourea, etc. The sulfanilyl derivatives **A14** and **B14** obtained as mentioned above were then chosen for further elaboration, with the goal of obtaining more efficient thrombin inhibitors. Two main approaches have been used for attaining this purpose: (1) Attachment of protected amino acyl/dipeptidyl moieties to the N-4 atom of the two leads **A14** and **B14**, based primarily on the well-known scaffold Phe-Pro-Arg,[3−5,7,48] which ensures effective binding within thrombin active site. Since our compounds already possessed the S1 anchoring moiety, basically Phe and Pro substitutions were performed initially. The terminal amino moiety of these amino acids was protected by means of the classical benzyloxycarbonyl group (Cbz) or by the new tosylureido (ts) group, which has not been used up to now for the design of serine protease inhibitors. Our choice is motivated by the following facts: the toluenesulfonyl group contained in this moiety might participate in interactions with the hydrophobic moieties of the thrombin binding site (such as S2 or/and S4; in the case of compounds possessing a larger molecule, see later in the text); additionally, the $SO_2NHCONH$ part of the molecule might also be critical for binding, as in the X-ray structure of the NAPAP adduct of thrombin it was observed that the SO_2NHCH_2CO moiety forms two strong hydrogen bonds with Gly 216.[49] Thus, the dipeptidyl derivatives ts-D-Phe-Pro- as well as Cbz-D-Phe-Pro- of the two leads have been prepared (compounds **A,B(36**,**37)**), and in fact they proved to be the most potent inhibitors in the whole series. (2) Considering the fact that the well-known serine protease substrate Chromozym TH (Ts-Gly-Pro-Arg-*p*-nitroanilide,Ts = 4-toluenesulfonyl; it should not be confused with "ts" which is our abbreviation for tosylureido) possesses a strong affinity for thrombin (K_m between 4 and 12 μM)[23] we decided to use part of this scaffold for designing sulfanilylguanidine- and sulfanilyl-*O*-methylisourea-based inhibitors. This scaffold was modified as follows: the tosyl group has been changed to the ts (tosylureido) one; the amino terminal Gly was either maintained (such as in **A38**) or changed to β-Ala (such as in **A39**) in order to study the influence of the length of this part of the molecule to binding to the enzyme. Most importantly, the Pro of the original scaffold was changed to the much more polar His in the new compounds **A38** and **A39**. All these changes led to effective thrombin inhibitors in some cases. Finally, among the different other dipeptides attached to the two lead molecules mentioned above (data not shown), the one leading to strong and selective thrombin inhibitors was that based on the ts-Pro-Gly scaffold, in derivatives **A40** and **B40**. No structural data are available up to now for explaining the high affinity of such inhibitors to thrombin.

Serine Protease Inhibition. Data of Tables 1 and 3 show that the sulfonylguanidines **A1**−**A32** and the corresponding sulfonyl-*O*-methylisoureas **B1**−**B32** possess weak thrombin inhibitory properties, comparable with those of the lead benzamidine **9**. The main difference between the two classes of inhibitors regards the decreased affinity of our compounds for trypsin, as compared to benzamidine. These data also clearly show that the presence of an aromatic/heterocyclic ring into



B34



A39

the molecule of such a compound highly enhances its affinity for thrombin. The aliphatic compounds, such as the dimethylamino derivatives **A1** and **B1** or the perfluoroalkylsulfonylguanidines **A3**, **A29**, and **A30**, and the corresponding perfluoroalkylsulfonyl-*O*-methylisoureas **B3**, **B29**, and **B30**, were among the most ineffective thrombin/trypsin inhibitors in the whole series of prepared compounds, with lower affinities than benzamidine **9** for the first enzyme. The derivatives of the substituted-benzenesulfonylguanidine/*O*-methylisourea on the other hand possessed generally a higher affinity for thrombin, as compared to benzamidine. Typically, they inhibited thrombin with K_I's between 100 and 300 nM but possessed a lower affinity for trypsin as compared to benzamidine, which is a highly desirable feature for an inhibitor to be developed for clinical applications. Moieties leading to effective thrombin inhibitors were: *p*-halogenophenyl, *p*-nitrophenyl, 3-chloro-4-nitrophenyl, *p*- and *m*-aminophenyl, perfluorophenyl, *p*-carboxyphenyl, 4-methoxy-3-aminophenyl, 1- and 2-naphthyl, or 2-thienyl, among others. The sulfonylguanidines of type **A** were slightly more active as thrombin/trypsin inhibitors as compared to the corresponding sulfonyl-*O*-methylisoureas **B**. In the two series of obtained compounds, the best thrombin inhibitors proved to be those containing *p*- or *m*-amino moieties. Thus, one such compound in each series has been chosen as a lead for further elaboration, with the scope of obtaining more efficient inhibitors. The two selected compounds were the sulfanilyl derivatives **A14** and **B14**, respectively. In the two subseries of amino acyl/dipeptidyl-containing inhibitors **A,B(33−40)**, the following facts should be noted (Tables 2 and 4): (i) All these compounds behave as stronger thrombin inhibitors (K_I's in the range of 12−62 nM) as compared to the two leads from which they were obtained, whereas their affinity for trypsin remained relatively low (K_I's in the range of 1200−1540 nM). Furthermore, the affinity of the best inhibitors for thrombin is comparable to or slightly better than that of the clinically used argatroban and inogatran, being at the same time less effective than NAPAP. On the other hand, the thrombin over trypsin selectivity of our compounds is much better as compared to that of other reported inhibitors. (ii) Protected dipeptide derivatives (such as **A**,**B**(**36**−**40**)) were more effective inhibitors as compared to the protected amino acyl derivatives **A**,**B**(**33**−**35**). (iii) The ts moiety seems to be slightly more beneficial than the Cbz one for potency as thrombin inhibitors of the obtained compounds. (iv) Prolongation of the molecule from Gly to *β*-Ala in compounds **A39** and **B39** is also beneficial for their potency/selectivity profiles. (v) Amino acyl/dipeptidylsulfaguanidines **A** were more effective thrombin inhibitors as compared to the corresponding acyl/dipeptidyl-*O*-methylsulfonylisoureas **B**.

Although the thrombin and trypsin S1 pockets are quite similar, the substitution Ser190Ala has as a consequence in the fact that a hydrogen-bonding partner for the P1 residue (OH from Ser 190 in thrombin) is lost in trypsin.[50] The specificity for thrombin over trypsin of the compounds reported here might be thus explained by considering the characteristic features of the thrombin active site, which, unlike that of trypsin, is also partially occluded.[51] Furthermore, thrombin shows a greater selectivity for the P2 and P3 residues, as compared to trypsin,[52] and it has also been showed that it is the residue Glu 192 that plays a major role in the determination of the specificity of the coagulation proteases toward their substrates/inhibitors.[53] All these factors might thus explain the good selectivity profile of the compounds reported here for thrombin over trypsin. X-ray crystallographic work is in progress in our laboratory in order to better understand the binding mode to the serine proteases of these new classes of inhibitors.

QSAR Calculations. 1. Guanidines. 1.1. Compounds Without Peptide Chains. 1.1.1. Thrombin. The best single-variable equation involved Δ_H, with $R^2 = 0.47$. Of the best 10 equations on the C_p criterion, the only one with all terms statistically significant was:

$$\log K_I = C_1\mu_x + C_2\Delta_H + C_3\Delta_L + C_4\Pi + C_5 \quad (1)$$

i	*1*	*2*	*3*	*4*	*5*
C_i	-0.0163	-0.233	0.190	-0.00275	2.55
σ	0.0038	0.047	0.077	0.00072	0.09
α	.00038	.00008	.02294	.00109	.00000

$N = 25$, $R^2 = 0.752$, $Q^2 = 0.626$, $s = 0.083$, $\Lambda = 1.46$, $F = 15.17$, $P = 7 \times 10^{-6}$

Here N is the number of drugs under consideration, R^2 is the squared correlation coefficient, Q^2 is the same based on the predicted residuals (the leave-one-out technique), s is the standard error of estimate, F is the Fisher variance ratio for the regression, and P is the corresponding probability (the statistical significance). The statistic Λ has been described above. For each variable in the regression, C_i is the regression coefficient, σ is the standard error, and α is the statistical significance.

One of Q_C, Q_{Nb}, or E_H entered next but was not significant by *t*-test, followed by either $\cos 2\Theta_H$ or $\sin 2\Theta_H$.

1.1.2. Trypsin. The best single-variable equation involved Δ_H. The best equation on the C_p criterion was:

$$\log K_I = C_1\Delta_H + C_2\Delta_L + C_3 \quad (2)$$

i	*1*	*2*	*3*
C_i	-0.0901	-0.110	3.225
σ	0.0332	0.046	0.042
α	0.01279	0.02490	0.00000

$N = 25$, $R^2 = 0.310$, $Q^2 = 0.153$, $s = 0.060$, $\Lambda = 1.13$, $F = 4.94$, $P = 0.016$

An equation with $R^2 = 0.393$ and almost equally good in C_p involved log P, E_L, and cos $2\Theta_H$ − sin $2\Theta_H$.

1.2. All Compounds. 1.2.1. Thrombin. The best equation on the C_p criterion was:

$$\log K_I = C_1\Pi_{zz} + C_2\Delta H_S + C_3 S^E{}_H + C_4 E_H + C_5 \quad (3)$$

i	*1*	*2*	*3*	*4*	*5*
C_i	-0.00382	0.00746	-257.8	-0.289	-9.93
σ	0.00033	0.00115	55.7	0.046	2.30
α	0.00000	0.00000	0.00008	0.00000	0.00018

$N = 33$, $R^2 = 0.943$, $Q^2 = 0.924$, $s = 0.106$, $\Lambda = 1.52$, $F = 116.8$, $P = 5 \times 10^{-17}$

A_w could substitute for Π_{zz} and Δ_H and E_L could replace E_H with little loss of fit.

1.2.2. Trypsin. The best equation found on the C_p criterion was:

$$\log K_I = C_1 S^N{}_S + C_2 S^N{}_{Na} + C_3 S^N{}_H + C_4\Delta_H + C_5 \cos 2\Theta_H + C_6 E_H + C_7 \quad (4)$$

i	*1*	*2*	*3*	*4*	*5*	*6*	*7*
C_i	-0.0586	-0.693	9.68	-0.151	0.0683	0.0666	1.425
σ	0.0150	0.244	1.60	0.035	0.016	0.0273	0.496
α	0.00057	0.00870	0.00000	0.00024	0.00019	0.02193	0.00792

$N = 33$, $R^2 = 0.673$, $Q^2 = 0.474$, $s = 0.040$, $\Lambda = 4.04$, $F = 8.93$, $P = 2 \times 10^{-5}$

2. Aminoguanidines. 2.1. Compounds Without Peptide Chains. 2.1.1. Thrombin. The only equation among the best 10 with all terms statistically significant was:

$$\log K_I = C_1\mu_x + C_2 E_H + C_3 \quad (5)$$

i	*1*	*2*	*3*
C_i	-0.0143	-0.246	-0.268
σ	0.0032	0.037	.371
α	0.00029	0.00000	.47777

$N = 25$, $R^2 = 0.693$, $Q^2 = 0.622$, $s = 0.090$, $\Lambda = 1.12$, $F = 24.85$, $P = 2 \times 10^{-6}$

The next variables to enter were Π or cos $2\Theta_L$, but neither was statistically significant.

2.1.2. Trypsin. The best equation found on the basis of C_p was the one-variable expression:

$$\log K_I = C_1 \sin 2\Theta_L + C_2 \quad (6)$$

$C_1 = 0.147 \pm 0.049$, $C_2 = 3.220 \pm 0.040$, $R^2 = 0.278$, $Q^2 = 0.146$, and $s = 0.058$

No other equation among the best 10 had all terms significant.

2.2. All Compounds. 2.2.1. Thrombin. The best equation on the basis of C_p had 10 terms and an R^2 of 0.956 and had two terms which did not achieve statistical significance. The best equation for which all terms were statistically significant was:

$$\log K_I = C_1\mu + C_2 Q_C + C_3 Q_{Nb} + C_4 S^N{}_H + C_5 \log P + C_6 E_{H+L} + C_7\Pi + C_8 \quad (7)$$

i	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
C_I	-0.0230	-13.7	-15.7	57.6	-0.0722	-0.159	-0.00464	-25.91
σ	0.0085	3.3	5.7	13.4	0.0247	.048	0.00038	7.65
α	.01202	.00036	.01077	.00024	.00717	.00311	0.00000	.00233

$N = 33$, $R^2 = 0.937$, $Q^2 = 0.895$, $s = 0.124$, $\Lambda = 4.72$, $F = 53.37$, $P = 2 \times 10^{-13}$

2.2.2. Trypsin. The best equation on the basis of C_p was:

$$\log K_I = C_1\mu_x + C_2\Delta H_S + C_3 \sin 2\Theta_L + C_4 \quad (8)$$

i	*1*	*2*	*3*	*4*
C_i	-0.00347	-0.000981	0.1549	3.169
σ	0.00136	0.000296	0.0098	0.036
α	0.01600	0.00250	0.00102	0.00000

$N = 33$, $R^2 = 0.463$, $Q^2 = 0.297$, $s = 0.049$, $\Lambda = 1.05$, $F = 8.33$, $P = 4 \times 10^{-4}$

3. Isoureas. 3.1. Compounds Without Peptide Chains. 3.1.1. Thrombin. The best equation in terms of C_p was:

$$\log K_I = C_1\mu_x + C_2\mu_z + C_3 Q_C + C_4 \log P + C_5\Delta_L + C_6 E_H + C_7\Pi + C_8 \quad (9)$$

i	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
C_i	-0.0169	-0.0412	-11.16	0.0829	0.194	-0.262	-0.00419	6.690
σ	0.0037	0.0143	2.58	0.0255	-.082	0.047	0.00086	1.519
α	.00027	.01046	.00046	.00474	.03012	.00003	0.00015	.00039

$N = 25$, $R^2 = 0.863$, $Q^2 = 0.677$, $s = 0.074$, $\Lambda = 2.14$, $F = 15.32$, $P = 8 \times 10^{-6}$

3.1.2. Trypsin. The best equation in terms of C_p was statistically nonsignificant:

$$\log K_I = C_1 Q_C + C_2 \sin 2\Theta_L + C_3 \quad (10)$$

$\log K_I = C_1 Q_C + C_2 \sin 2\Theta_L + C_3$

i	*1*	*2*	*3*
C_i	-3.526	0.0730	5.201
σ	1.881	0.0363	1.069
α	0.07425	0.05676	0.00007

$N = 25$, $R^2 = 0.210$, $Q^2 = -0.627$, $s = 0.066$, $\Lambda = 1.09$, $F = 2.91$, $P = 7 \times 10^{-2}$

3.2. All Compounds. 3.2.1. Thrombin. The best equation by the C_p criterion was:

$$\log K_I = C_1\mu_y + C_2 S^N{}_S + C_3 S^E{}_{Na} + C_4 \log P + C_5\Delta_H + C_6\Delta_L + C_7\Pi + C_8 \quad (11)$$

i	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
C_i	0.0107	-0.184	69.4	0.0914	-0.231	0.319	-0.00455	24.526
σ	0.0049	0.038	13.3	0.0195	0.056	0.090	0.00024	4.150
α	.03686	.00005	.00002	.00008	.00038	.00161	.00000	.00000

$N = 33$, $R^2 = 0.958$, $Q^2 = 0.920$, $s = 0.092$, $\Lambda = 2.04$, $F = 82.5$, $P = 1 \times 10^{-15}$

3.2.2. Trypsin. The best equation in terms of C_p was statistically nonsignificant:

$$\log K_I = C_1 \cos 2\Theta_L + C_2 E_L + C_3 \quad (12)$$

i	*1*	*2*	*3*
C_i	0.0815	0.0789	3.289
σ	0.0344	0.0381	0.050
α	0.02097	0.04696	0.00000

$N = 33$, $R^2 = 0.193$, $Q^2 = -0.681$, $s = 0.058$, $\Lambda = 1.23$, $F = 3.59$, $P = 4 \times 10^{-2}$

4. All Three Carbimido Derivatives. 4.1. Compounds Without Peptide Chains. 4.1.1. Thrombin. The best equation by the C_p criterion with all variables statistically significant was:

$$\log K_I = C_1 I_B + C_2\mu_x + C_3 S^E{}_O + C_4\Delta H_S + C_5 E_H + C_6 S^E{}_H + C_7 \quad (13)$$

i	*1*	*2*	*3*	*4*	*5*	*6*	*7*
C_i	0.165	-0.0117	-9.45	0.00145	-0.312	-54.3	-10.589
σ	0.039	-0.0021	1.73	0.00049	0.024	13.54	1.99
α	.00008	.00000	.00000	.00413	.00000	.00015	.00000

$N = 75$, $R^2 = 0.749$, $Q^2 = 0.698$, $s = 0.085$, $\Lambda = 4.31$, $F = 33.76$, $P = 1 \times 10^{-18}$

Figure 1 shows the fit of the data to this equation.

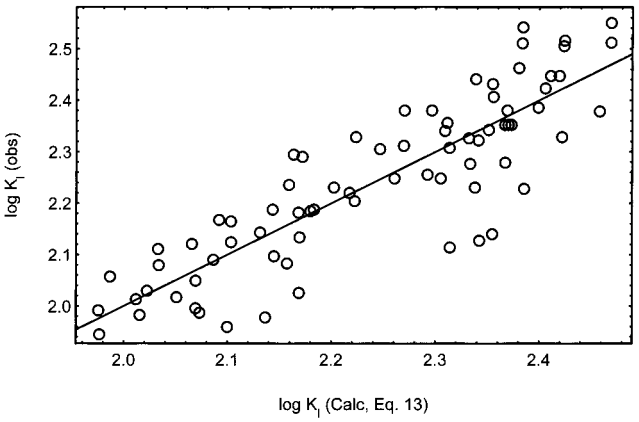


Figure 1. Fitted versus experimental log K_I for all nonpeptide-bearing compounds (75 compounds).

4.1.2. Trypsin. The best equations by the C_p criterion with all variables statistically significant were:

$$\log K_I = C_1 \log P + C_2\mu + C_3 E_L + C_4 S^N{}_{Na} + C_5 S^N{}_H + C_6 \sin 2\Theta_L + C_7 \quad (14)$$

i	*1*	*2*	*3*	*4*	*5*	*6*	*7*
C_i	0.0198	0.00440	0.112	-0.584	5.25	0.119	2.147
σ	0.0084	0.00203	0.022	0.213	1.45	0.023	0.292
α	.02175	.03386	.00001	0.00773	.00054	.00000	0.00000

$N = 75$, $R^2 = 0.493$, $Q^2 = 0.313$, $s = 0.055$, $\Lambda = 4.15$, $F = 11.02$, $P = 2 \times 10^{-8}$

and

$$\log K_I = C_1 I_B + C_2 \log P + C_3 E_L + C_4 \cos 2\Theta_L + C_5 \sin 2\Theta_L + C_6 \quad (15)$$

i	*1*	*2*	*3*	*4*	*5*	*6*
C_i	0.0405	0.0214	0.0983	-0.0257	0.1162	3.270
σ	0.0143	0.0087	0.0233	0.0223	0.0245	0.038
α	.00613	.01677	.00008	.25423	.00001	.00000

$N = 75$, $R^2 = 0.428$, $Q^2 = 0.242$, $s = 0.058$, $\Lambda = 1.26$, $F = 10.33$, $P = 2 \times 10^{-7}$

4.2. All Compounds. 4.2.1. Thrombin. The best equation by the C_p criterion with all variables statistically significant was:

$$\log K_I = C_1 \log P + C_2 I_B + C_3\mu_x + C_4\mu_z + C_5\Pi + C_6\Delta_L + C_7\Delta_H + C_8 \quad (16)$$

i	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
C_i	0.0755	0.0819	-0.0103	-0.0151	-0.00532	0.166	-0.201	2.981
σ	0.0133	0.0260	0.0025	-.0037	0.00020	0.046	0.038	0.120
α	.00000	.00221	.00011	.00012	.00000	.00050	.00000	.00000

$$N = 99,\ R^2 = 0.924,\ Q^2 = 0.908,\ s = 0.120,$$
$$\Lambda = 1.38,\ F = 157.3,\ P = 5 \times 10^{-48}$$

Figure 2 shows the fit of the data to this equation. When stepwise regression was carried out on the full data set with random reassignment of the dependent variable, allowing nine variables to enter, a statistical significance of 2×10^{-30} was obtained.[40,41]

4.2.2. Trypsin. The best equation by the C_p criterion with all variables statistically significant was:

$$\log K_I = C_1\mu_x + C_2 I_B + C_3\Delta H_S + C_4\Delta_H + C_5 \cos 2\Theta_H + C_6 \sin 2\Theta_L + C_7 Q_C + C_8 S^E{}_O + C_9 \quad (17)$$

i	1	2	3	4	5	6	7	8	9
C_i	-0.00375	0.0787	-0.000664	-0.0314	0.0267	0.0869	-0.359	-4.45	-0.300
σ	.00096	0.0159	0.000207	0.0157	0.0117	0.0215	0.095	1.17	0.908
α	.00018	.00000	.00187	.04778	.02492	.00011	.00026	.00027	.73179

$$N = 99,\ R^2 = 0.430,\ Q^2 = 0.243,\ s = 0.052,$$
$$\Lambda = 3.74,\ F = 8.45,\ P = 2 \times 10^{-8}$$

When the random reassignment technique was applied to this data set allowing eight descriptors to enter, a statistical significance of 0.001 was obtained for the stepwise procedure.

It is clear from the activity data that the compounds fall into three distinct groups. Least active are the aliphatic sulfonamide derivatives, for which the QSARs will not be considered. The aromatics are more active, probably because of favorable $\pi-\pi$ interactions. The aromatics bearing peptide chains are especially active in the case of thrombin inhibition. There is much greater variation among the thrombin inhibitory activity than the trypsin, and this is reflected in the better R^2 for the thrombin data. Although R^2 for trypsin inhibitory activity is poor, the standard errors of estimate for this activity tend to be very small and reflect a fit in which the lack of fit is little greater than the experimental error in the data.

The statistics of chance correlation in the stepwise regression has been calculated only for the pooled data sets, using all of the aromatics, and found to be of extremely high significance for the thrombin results and acceptable for the trypsin results. It would be hard to compute for the small data sets as variables had to be arbitrarily excluded from the calculation in order to get any results from the all possible subsets procedure. That is, there were too many variables for the number of cases in these subsets. This suggests that the results for the pooled data are much to be preferred.

In the equations for antithrombin activity, a term representing size such as A_w or Π or one of its components was nearly always present and of very high statistical significance, especially when the highly active peptide derivatives were included. This was not the case in equations for antitrypsin activity. The regressions on the pooled data from all three sets of compounds are of much higher significance and are more informative than that for the individual sets. For the aromatics with peptide appendages, the FOPA descriptors all have nearly the same value, resulting in many nearly coincident points in the parameter space. It is perhaps more informative to consider just the 75 compounds without such groups.

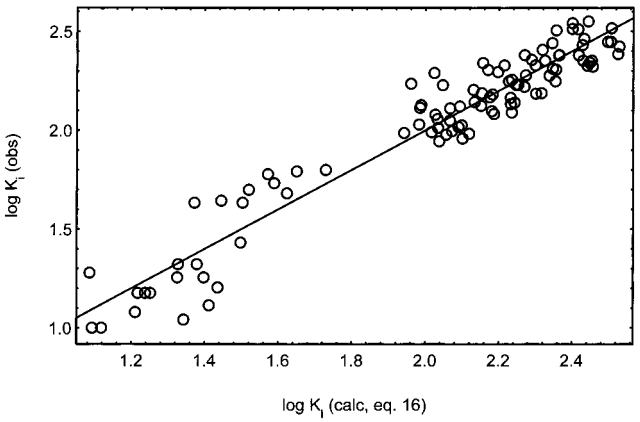


Figure 2. Fitted versus experimental log K_I for all compounds (99 compounds).

The dependence on polarizability or size of antithrombin activity becomes greater when the compounds with bulky peptide substituents, which are highly active, are included but is nevertheless present in the other compounds. In this data set, as is usually the case, polarizability correlates strongly with size, and both can explain activity, the former through dispersion forces and the latter through steric effects. For trypsin inhibition, the FOPA variables, orbital energies, and orbital energy splitting are the most important. E_H tends to be more important for thrombin inhibition and E_L for trypsin inhibition. The physical significance of some of these correlations becomes apparent when it is realized that Θ_H and Θ_L are simply related to the angles which the nodes in the HOMO and LUMO of the drug molecule make at the center of the benzene ring with atom 1.[54] The angle the node of the highest occupied π orbital (HOPO) makes with the reference atom (here the atom of the benzene ring to which the sulfonamide S is bound) is $\Theta_H + \pi/6$. The two lowest unoccupied π orbital (LUPO) nodes, mutually perpendicular at the center of the ring in the approximation, make angles of $^1/_2(\Theta_H - \pi/6)$ and $^1/_2(\Theta_H + 5\pi/6)$. Part of the binding force between the enzyme and inhibitor is due to an interaction between a frontier π orbital on the drug with a π system on some component of the enzyme, and this is greatest when the splitting of the frontier orbital of benzene is greatest and when the directions of the nodes in the drug orbital most nearly correspond with those on the enzyme.

A drug bound to a receptor or enzyme is severely restricted in the orientations it can assume. If an aromatic group in such a drug is interacting with an aromatic group on the receptor, the interaction will be strongest when the nodes in the interacting orbitals, usually the HOPO and LUPO, most nearly coincide when the drug is in the bound orientation. Thus for a member of a congeneric series of such drugs, there will be an optimum Θ_H and/or Θ_L, at which the interaction will be maximal. From eq 16 (all compounds without peptide chains), trypsin inhibitory activity is maximal when Θ_L is 144.6°. This corresponds to the LUPO node making an angle at the center of the benzene ring of 3.1° with the carbon atom bearing the sulfonamide

group. Similarly, from eq 17 (all compounds) a maximum activity occurs when the node angle for the LUMO is 75° and that for the HOPO is 120°. These latter angles are not most probable values, as statistically insignificant terms have been discarded.

The difference between E_H and E_L is an approximation to the Pearson hardness of the molecule, and it is well-established that hard−hard and soft−soft interactions are preferred over hard−soft. By Koopmans' theorem, Δ_H and Δ_L are related to the hardness of the corresponding radical cation and radical anion, respectively, and from the origins of the FOPA descriptors, it may be supposed that the smaller Δ_H and Δ_L, then the more easily the contributing orbitals on one molecule may be distorted to allow a match of nodes with the corresponding orbitals on the other. When the Δ_H and Δ_L terms are small however, if Θ_H or Θ_L is favorable there is an unfavorably oriented orbital lying just below the frontier orbital. In this circumstance the orientation effect may be small.

From regressions in which statistically nonsignificant terms were retained, and using the theoretical interpretation of Θ_H and Θ_L, maximum antitrypsin activity of the drug occurs when the HOMO node makes an angle of 97.5° with the sulfonamide group at the center of the benzene ring and one of the LUMO nodes makes an angle of 3.0° with the same group. For antithrombin activity, the corresponding angles are 120.8° for the HOMO and 77.8° for the LUMO, but these must be regarded as very doubtful, because of the very poor statistical significance of the FOPA variables for antithrombin activity. The results indicate that the benzene moiety of the drugs interacts with a π electron system in the enzyme and that this interaction is much more pronounced in trypsin than in thrombin. Dependence of activity on the FOPA descriptors is perhaps the clearest indication yet of a quantum effect in pharmacology.

The known structure of thrombin lends plausibility to the hypothesis that aromatic residues of the enzyme bond to the drug. The two insertion loops around residues Trp 60D and Trp 148 in thrombin restrict the active site region from opposite sides, producing a particularly narrow active site, as compared to that of other serine proteases. Other aromatic residues within the active site, such as Tyr 60A, Trp 215, Tyr 228, or Phe 227, were also shown by X-ray crystallography to interact with inhibitors/substrates.

It is expected that the FOPA parameters, Θ_H and Θ_L, together with the associated Δ_H and Δ_L will become a powerful tool in the design of drugs based on multisubstituted benzene whenever the benzene ring is involved in binding to the receptor. The maximal interaction will occur when the nodes of the interacting orbitals coincide. This will depend on Θ_H and Θ_L for the drug, the similar terms for the receptor, and the relative orientations of the two in the bound complex. Ideally, the HOPO of the benzene ring of the drug interacts with the HOPO orbital of similar symmetry on part of the receptor, giving in the complex two new orbitals, one of which is higher in energy than either of the originals and the other lower. A similar interaction occurs with the LUPO orbitals, resulting in a reduced HOPO−LUPO gap for the complex. In cases where there is more than one set

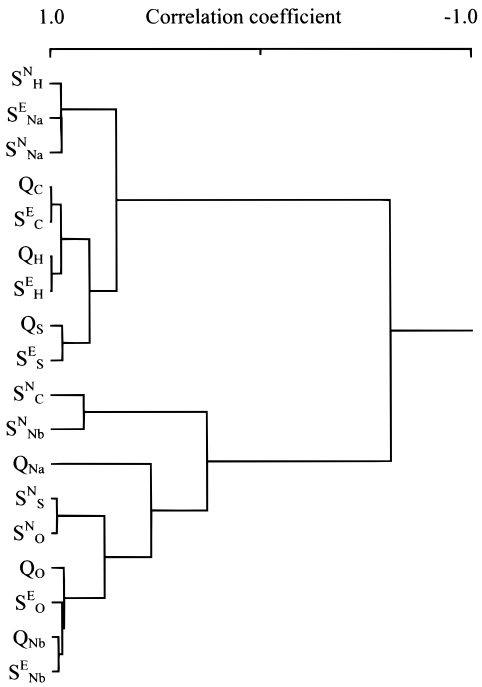


Figure 3. Correlation dendrogram of the sulfonylguanidine/isourea/aminoguanidine electronic descriptors for all compounds.

of π orbitals, it should of course be confirmed that the appropriate orbitals are being used. This can easily be seen from a printout of the eigenvectors, with an appropriate molecular orientation. The effective component of the receptor is probably a π electron system on an amino acid residue, such as benzene, indole, imidazole, or guanidine.

The superdelocalizabilities S^E and S^N are measures of susceptibility of the atom in question to electrophilic and nucleophilic attack. It should be noted that these terms and the Mulliken charges are strongly intercorrelated. A correlation dendrogram for them is shown in Figure 3, for the 99 compounds. While S^E correlates positively with Q, the expected negative correlation with S^N is not always apparent. Caution is needed in the interpretation of S^N as it is numerically unstable, with the small and not very accurately known virtual orbital energies appearing in denominators, leading to large errors. This is similar to but less severe than the problem encountered by Gomez-Jeria with CNDO calculations.[55] They become rather dependent to the local conformation in the neighborhood of the atoms in question. Charges tend to relate to fast reactions where the transition state resembles the reactants, while superdelocalizabilities are more relevant when they differ.

Since the dependent variable is log K_I, a negative regression coefficient indicates that activity increases with increasing value of the correspoding descriptor. Examination of eqs 1−16 shows that large Δ_H favors high activity for both trypsin and thrombin. Large Δ_L increases activity against trypsin but reduces antithrombin activity. Antithrombin activity is disfavored and antitrypsin activity is favored by high solvation energy, as calculated by the COSMO procedure. Both are favored by large x and z components of the dipole moment and by low lipophilicity.

The cross-validated residuals corresponding to eq 16 were examined more closely. Those larger than 0.1 corresponded to **A8**, **A10**, **A11**, **B4**, **B10**, **B11**, **C8,** and **C11**. None of these reached 0.2. They do not correspond to any particular structure except nitro, and in particular, none correspond to an anion. The anions are well-described by eq 16, and there seems to be no significant lack of fit.

Experimental Section

General. Melting points: heating plate microscope; not corrected. IR spectra: KBr pellets; 400−4000 cm^{-1} Perkin-Elmer 16PC FTIR spectrometer. ^1H NMR spectra: Varian 300CXP apparatus; chemical shifts are expressed as δ values relative to Me_4Si as standard; J values are in Hz. Elemental analysis: ±0.4% of the theoretical values, calculated for the proposed formulas, data not shown; Carlo Erba Instrument CHNS elemental analyzer, model 1106. All reactions were monitored by thin-layer chromatography (TLC) using 0.25-mm precoated silica gel plates (E. Merck). Preparative HPLC was performed on a Dynamax-60A column (25 × 250 mm), with a Beckman EM-1760 instrument. The detection wavelength was 254 nm. Sulfonyl halides, guanidine, *O*-methylisourea hydrochloride, triethylamine, carbodiimides, amino acids, Cbz-amino acids, dipeptides, and tosyl isocyanate used in the syntheses were commercially available compounds (from Sigma, Acros or Aldrich). Acetonitrile, acetone, dioxane, ethyl acetate (E. Merck, Darmstadt, Germany) or other solvents used in the synthesis were doubly distilled and kept on molecular sieves in order to maintain them in anhydrous conditions. Inogatran was from Astra Hassle (Molndal, Sweden). Benzamidine, NAPAP, human thrombin, human trypsin and Chromozym TH were from Sigma Chemical Co. (St. Louis, MO).

General Procedure for the Preparation of Compounds A,B(1−32). Methods A and B: An amount of 10 mM sulfonyl halide (chloride for method A and fluoride for method B) was dissolved in 50 mL of acetone and the stoichiometric amount of guanidine/*O*-methylisourea·HCl dissolved in 5 mL of water was added dropwise, together with the stoichiometric amount of solid $NaHCO_3$ needed for the neutralization of the acid. The mixture was magnetically stirred at 25 °C for 5 h, then the reaction mixture was acidified with 0.1 N HCl solution till pH 3. The obtained sulfonyl derivatives generally precipitated by leaving the above-mentioned reaction mixture at 4 °C overnight and were then filtered and recrystallized from ethanol.

Method C: An amount of 5 mmol of guanidine/*O*-methylisourea·HCl and 0.84 mL (5 mmol) of triflic anhydride were suspended in 10 mL of acetone and 0.35 mL (5 mmol) of triethylamine was added dropwise. The mixture was magnetically stirred at 4 °C for 5 h. The solvent was then evaporated in vacuo and the tan residue treated with 5 mL of cold water. The triflate salt of triethylamine being water soluble was separated from the triflated guanidine/*O*-methylisourea (much less water soluble) by a simple filtration. The latter compounds were recrystallized from 2-propanol.

Method D: An amount of 10 mmol of guanidine/*O*-methylisourea·HCl, 0.70 mL (10 mmol) of triethylamine and 10 mmol of sulfobenzoic cyclic anhydride or tetrabromo-*O*-sulfobenzoic cyclic anhydride were heated at reflux in 50 mL of anhydrous acetonitrile for 2 h, with a small amount of *p*-toluenesulfonic acid added as catalyst. After evaporation of the solvent, the products were treated with 10 mL of water and the precipitated derivatives filtered and recrystallized from ethanol.

General Procedure for the Preparation of *N*-Tosylureido-Protected Amino Acids/Dipeptides. An amount of 10 mM of amino acid/dipeptide was suspended/dissolved in 50 mL of anhydrous acetone and 1.97 g (1.52 mL, 10 mM) of tosyl isocyanate was added dropwise. The reaction mixture was stirred at 4 °C for 4 h, when by means of TLC it was observed that the reaction was completed. Evaporation of the solvent in vacuo afforded white foams of *N*-tosylureido-protected amino acid derivatives, which were recrystallized from ethanol−water (1:1, v/v).

General Procedure for the Preparation of Compounds 8. An amount of 1 mM *N*-Cbz- or *N*-tosylureido-protected amino acid/dipeptide **7** was dissolved/suspended in 25 mL of anhydrous acetonitrile or acetone and then treated with 197 mg (1 mM) of sulfanilylcyanamide **6** and 190 mg (1 mM) of EDCI·HCl. The reaction mixture was magnetically stirred at room temperature for 15 min, then 30 μL (2mM) of triethylamine was added and stirring was continued for 16 h at 4 °C. The solvent was evaporated in vacuo and the residue taken up in ethyl acetate (5 mL), poured into a 5% solution of sodium bicarbonate (5 mL) and extracted with ethyl acetate. The combined organic layers were dried over sodium sulfate and filtered and the solvent removed in vacuo. Preparative HPLC (Dynamax-60A column (25 × 250 mm); 90% acetonitrile/8% methanol/2% water; flow rate of 30 mL/min) afforded the pure compounds **8** as colorless solids.

General Procedure for the Preparation of Compounds B33−B40. An amount of 2 mM derivative **8** was dissolved in 30 mL of absolute MeOH and 15 mL of absolute dioxane and cooled to 0 °C, and 0.45 g (12 mM) of dried gaseous HCl was introduced into the reaction mixture. The obtained solution was kept at 4 °C for 4 days, then poured into 100 mL of diethyl ether, and the resulting precipitate was triturated with 50 mL of THF until crystallization took place. The obtained crystals were washed with diethyl ether and dried in vacuo. Yields were almost quantitative (95−97%).

General Procedure for the Preparation of Compounds A33−A40. An amount of 1 mM derivative **B33−B40** was suspended in 30 mL of methanol and an excess of ethanolic ammonia solution was added. The mixture was heated on a steam bath at 60 °C for 4 h. Then the solvents were evaporated in vacuo and the residue retaken in 25 of mL of ethanol. Addition of ethyl acetate and 2 mL of 2 N HCl solution led to the precipitation of hydrochloride salts of compounds **A33−A40**. The neutral compounds were obtained by neutralization of the salts with solid sodium bicarbonate.

4-(4-Toluenesulfonylureido-*β*-alanyl-L-histidinylamido)-benzenesulfonylguanidine, A39: mp 279 °C; ^1H NMR (DMSO) δ 2.63 (s, 3H, C$H_3C_6H_4$), 2.79−2.88 (m, 2H, CH_2 of β-Ala), 3.11−3.26 (m, 2H, CH_2 of β-Ala), 3.34−3.45 (m, 2H, CHCH_2 of His), 4.57−4.63 (m, 1H, CHCH$_2$ of His), 7.32 (s, 1H, CH-5 of His), 7.56 (d, $^3J_{HH}$ = 8.1, 2H, H_{ortho} of $CH_3C_6H_4$), 7.69 (d, $^3J_{HH}$ = 7.9, 2H, H_{ortho} of $XNHC_6H_4SO_2$), 7.87 (d, $^3J_{HH}$ = 8.1, 2H, H_{meta} of $CH_3C_6H_4$), 7.94 (d, $^3J_{HH}$ = 7.9, 2H, H_{meta} of $XNHC_6H_4SO_2$), 8.35 (s, 1H, CH-2 of His), 9.21 (br s, 4H, SO_2N=C$(NH_2)_2$); ^{13}C NMR (DMSO) δ 25.9 (s, $CH_3C_6H_4$), 33.3 (s, CH_2 of His), 37.4 (s, NHCH_2CH_2 of β-Ala), 40.8 (s, CH_2CH_2CO of β-Ala), 59.6 (s, $CHCH_2$ of His), 122.2 (s, *C*-4 of His), 130.6 (s, C_{meta} of $CH_3C_6H_4$), 130.9 (s, C_{meta} of $XNHC_6H_4SO_2$), 131.8 (s, *C*-5 of His), 134.2 (s, C_{ortho} of $CH_3C_6H_4$), 135.1 (s, C_{ortho} of $XNHC_6H_4SO_2$), 137.2 (s, *C*-2 of His), 139.1 (s, C_{para} of $CH_3C_6H_4$), 140.4 (s, C_{para} of $XNHC_6H_4SO_2$), 145.6 (s, C_{ipso} of $CH_3C_6H_4$), 148.5 (s, C_{ipso} of $XNHC_6H_4SO_2$), 149.1 (s, NH*C*ONH), 175.6 (s, CH_2CO of β-Ala), 176.4 (s, *C*ONH of His), 178.9 (s, SO_2N=*C*). Anal. ($C_{24}H_{29}N_9O_7S_2$) C, H, N, S.

4-(4-Toluenesulfonylureido-D-phenylalanylamido)benzenesulfonyl-*O*-methylisourea, B34: mp 263 °C; ^1H NMR (DMSO) δ 2.65 (s, 3H, C$H_3C_6H_4$), 3.10−3.55 (m, 2H, CH_2CH of Phe), 3.76 (s, 3H, OCH_3), 4.08 (dd, $^3J_{HH}$ = 5.0, $^3J_{HH}$ = 7.8, 1H, CH_2CH of Phe), 7.29−7.58 (m, 7H, H_{ortho} of tosyl and H_{arom} of Phe), 7.65 (d, $^3J_{HH}$ = 7.9, 2H, H_{ortho} of $XNHC_6H_4SO_2$), 7.90 (d, $^3J_{HH}$ = 7.9, 2H, H_{meta} of $XNHC_6H_4SO_2$), 7.95 (d, $^3J_{HH}$ = 8.3, 2H, H_{meta} of tosyl), 9.13 (br s, 2H, SO_2N=C(NH_2)OMe); ^{13}C NMR (DMSO) δ 26.2 (s, $CH_3C_6H_4$), 41.7 (s, CH_2CH of Phe), 50.8 (s, OCH_3), 59.3 (s, $CHCH_2$ of Phe), 130.9 (s, C_{para} of Phe), 132.3 (s, C_{meta} of $CH_3C_6H_4$), 133.5 (s, C_{meta} of $XNHC_6H_4SO_2$), 133.8 (s, C_{meta} of Phe), 134.4 (s, C_{ortho} of Phe), 135.1(s, C_{ortho} of $CH_3C_6H_4$), 135.5 (s, C_{ortho} of $XNHC_6H_4SO_2$), 141.5 (s, C_{ipso} of Phe), 142.9 (s, C_{para} of $XNHC_6H_4SO_2$), 145.0 (s, C_{para} of $CH_3C_6H_4$), 148.6 (s, C_{ipso} of $CH_3C_6H_4$), 148.9 (s, C_{ipso} of

$XNHC_6H_4SO_2$), 149.1 (s, NH*C*ONH), 175.1 (s, $SO_2N{=}C$). Anal. ($C_{25}H_{27}N_5O_7S_2$, 573.65) C, H, N, S.

pK_a Determination. The half-neutralization point was measured by titrating the organic acids/bases with 0.05 N NaOH and 0.05 N HCl in EtOH−water (30%, v/v), using a glass electrode, as described by Bell and Roblin[56] for the structurally related antibacterial sulfonamides.

Enzyme Assays: K_I Determinations. Human thrombin and human trypsin were purchased from Sigma Chemical Co. (St. Louis, MO); their concentrations were determined from the absorbance at 280 nm and the extinction coefficients furnished by the supplier. The activity of such preparations was in the range of 2500−3000 NIH units/mg. The potency of standard and newly obtained inhibitors was determined from the inhibition of the enzymatic (amidolytic) activity of these serine proteases, at 21 °C, using Ts-Gly-Pro-Arg-pNA (Chromozym TH) from Sigma as substrate, by the method of Lottenberg et al.[23] The substrate was reconstituted as 4 mM stock in ultrapure water and brought to pH 4 with hydrochloric acid. Substrate concentrations were determined from absorbance at the isosbestic wavelength for the peptide-*p*-nitroanilide−*p*-nitroaniline mixtures. Extinction coefficients of 8270 $L{\cdot}mol^{-1}{\cdot}cm^{-1}$ in the used buffer (0.01 M Hepes−0.01 M Tris−0.1 M NaCl−0.1% poly(ethylene glycol) 6000) were employed. The rate of *p*-nitroanilide hydrolysis was determined from the change in absorbance at 405 nm using an extinction coefficient for *p*-nitroaniline of 9920 $L{\cdot}mol^{-1}{\cdot}cm^{-1}$ for the above-mentioned reaction buffer. Measurements were made using a Cart 3 spectrophotometer interfaced with a PC. Initial velocities were thus estimated using the direct linear plot-based procedure as reported by Lottenberg et al.[23] K_I's were then determined according to Dixon, using a linear regression program.[57] The K_I values determined are the means of at least three determinations.

Acknowledgment. This research was financed in part by the EU Grant ERB CIPDCT 940051.

Supporting Information Available: Tables S1−S6 giving the calculated descriptors for the 99 drugs considered in the QSAR section. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Shafer, J. A. Cardiovascular chemotherapy: anticoagulants. *Curr. Opin. Chem. Biol*. **1998**, *2*, 458−465.

(2) (a) Stürzebecher, J.; Meier, J. Synthetic low-molecular weight thrombin inhibitors: The ideal antithrombotic drug? *J. Enzyme Inhib*. **1995**, *9*, 1−2. (b) Stürzebecher, J.; Prasa, D.; Hauptmann, J., Vieweg, H.; Wikstrom, P. Synthesis and structure−activity relationships of potent thrombin inhibitors: Piperazides of 3-amidinophenylalanine. *J. Med. Chem*. **1997**, *40*, 3091−3099.

(3) (a) Pavone, V.; De Simone, G.; Nastri, F.; Galdiero, S.; Staiano, N.; Lombardi, A.; Pedone, C. Multiple binding mode of reversible synthetic thrombin inhibitors. A comparative structural analysis. *Biol. Chem*. **1998**, *379*, 987−1006. (b) De Simone, G.; Lombardi, A.; Galdiero, S.; Della Morte, R.; Staiano, N.; Pedone, C.; Bolognesi, M.; Pavone, V. Hirunorms are true hirudin mimetics: the crystal structure of human α-thrombin: hirunorm V complex. *Protein Sci*. **1998**, *7*, 243−253.

(4) (a) Stubbs, M. T.; Bode, W. A player of many parts: the spotlight falls on thrombin's structure. *Thromb. Res*. **1993**, *69*, 1−58. (b) Engh, R. A.; Brandstetter, H.; Sucher, G.; Eichinger, A.; Baumann, U.; Bode, W.; Huber, R.; Poll, T.; Rudolph, R.; von der Saal, W. Enzyme flexibility and "weak" interactions characterize thrombin-ligand interactions: implications for drug design. *Structure* **1996**, *4*, 1353−1362.

(5) (a) Babine, R. E.; Bender, S. L. Molecular recognition of protein−ligand complexes: Application to drug design. *Chem. Rev*. **1997**, *97*, 1359−1472. (b) Salemme, F. R.; Spurlino, J.; Bone, R. Serendipity meets precision: the integration of structure-based drug design and combinatorial chemistry for efficient drug discovery. *Structure* **1997**, *5*, 319−324.

(6) Eriksson, U. G.; Renberg, L.; Bredberg, U.; Teger-Nilsson, A. C.; Regardh, C. G. Animal pharmacokinetics of inogatran, a low-molecular-weight thrombin inhibitor with potential use as an antithrombotic drug. *Biopharm. Drug Dispos*. **1998**, *19*, 55−64.

(7) (a) Bajusz, S.; Barbas, E.; Tolnay, P.; Szell, E.; Bagdy, D. Inhibition of thrombin and trypsin by tripeptide aldehydes. *Int. J. Pept. Protein Res*. **1978**, *12*, 217−221. (b) Bajusz, S.; Szell, E.; Bagdy, D.; Barbas, E.; Horvath, G.; Dioszegi, M.; Fittler, Z.; Szabo, G.; Juhasz, A.; Tomori, E.; Szilagyi, G. Highly active and selective anticoagulants. D-Phe-Pro-Arg-H, a free tripeptide aldehyde prone to spontaneous inactivation, and its stable *N*-methyl derivative D-Me-Phe-Pro-Arg-H. *J. Med. Chem*. **1990**, *33*, 1729−1735.

(8) Krishnan, R.; Zhang, E.; Hakansson, K.; Arni, R. K.; Tulinsky, A.; Lim-Wilby, M. S.; Levy, O. E., Semple, J. E., Brunck, T. K. Highly selective mechanism-based thrombin inhibitors: structures of thrombin and trypsin inhibited with rigid peptidyl aldehydes. *Biochemistry* **1998**, *37*, 12094−13103.

(9) Claeson, G.; Philipp, M.; Agner, E.; Scully, M. F.; Metternich, R.; Kakkar, V. V.; DeSoyza, T.; Niu, L. H. Benzyloxycarbonyl-D-Phe-Pro-methoxypropylboroglycine: a novel inhibitor of thrombin with high selectivity containing a neutral side chain at the P1 position. *Biochem. J*. **1993**, *290*, 309−312.

(10) (a) Stürzebecher, J.; Vieweg, H.; Wikstrom, P.; Turk, D.; Bode, W., Interaction of thrombin with benzamidine-based inhibitors. *Biol. Chem. Hoppe-Seyler* **1992**, *373*, 491−496. (b) Stürzebecher, J.; Prasa, D.; Bretschneider, E.; Bode, W.; Bauer, M.; Brandstetter, H.; Wikstrom, P.; Vieweg, H. New developments in the field of benzamidine-derived thrombin inhibitors. In *DIC − Pathogenesis, Diagnosis, and Therapy of Disseminated Intravascular Fibrin Formation*; Muller-Berghaus, G., Madlener, K., Blomback, M., ten Cate, J. W., Eds.; Excerpta Medica: Amsterdam, London, New York, Tokyo, 1993; pp 183−190.

(11) (a) Adang, A. E.; Lucas, H.; de Man, A. P.; Engh, R. A.; Grootenhuis, P. D. Novel acylguanidine containing thrombin inhibitors with reduced basicity at the P1 moiety. *Bioorg. Med. Chem. Lett*. **1998**, *8*, 3603−3608. (b) Lumma, W. C.; Witherup, K. M.; Tucker, T. J.; Brady, S. F.; Sisko, J. T.; Naylor-Olsen, A. M.; Lewis, S. D.; Lucas. B. J.; Vacca, J. P. Design of novel, potent, noncovalent inhibitors of thrombin with nonbasic P-1 substructures: rapid structure−activity studies by solid-phase synthesis. *J. Med. Chem*. **1998**, *41*, 1011−1013. (c) Reers, M.; Koschinsky, R.; Dickneite, G.; Hoffmann, D.; Czech, J.; Stuber, W. Synthesis and characterisation of novel thrombin inhibitors based on 4-amidinophenylalanine. *J. Enzyme Inhib*. **1995**, *9*, 61−72.

(12) (a) Levy, O. E.; Semple, J. E.; Lim, M. L.; Reiner, J.; Rote, W. E.; Dempsey, E.; Richard, B. M.; Zhang, E.; Tulinsky, A.; Ripka, W. C.; Nutt, R. F. Potent and selective thrombin inhibitors incorporating the constrained arginine mimic L-3-piperidyl(N-guanidino)alanine at P_1. *J. Med. Chem*. **1996**, *39*, 4527−4530. (b) Semple, J. E.; Rowley, D. C.; Brunck, T. K.; Ha-Uong, T.; Minami, N. K.; Owens, T. D.; Tamura, S. Y.; Goldman, E. A.; Siev, D. V.; Ardecky, R. J.; Carpenter, S. H.; Ge, Y.; Richard, B. M.; Nolan, T. G.; Hakanson, K.; Tulinsky, A.; Nutt, R. F.; Ripka, W. C. Design, synthesis, and evolution of a novel selective, and orally bioavailable class of thrombin inhibitors: P1-argininal derivatives incorporating P3−P4 lactam sulfonamide moieties. *J. Med. Chem*. **1996**, *39*, 4531−4536.

(13) Sixma, J. J.; de Groot, P. G. The ideal antithrombotic drug. *Thromb. Res*. **1992**, *68*, 507−512.

(14) Okamoto, S.; Kinjo, K.; Hijikata, A.; Kikumoto, R.; Tamao, Y.; Ohkubo, K.; Tonomura, S. Thrombin inhibitors. 1. Ester derivatives of Nα-(arylsulfonyl)-L-arginine. *J. Med. Chem*. **1980**, *23*, 827−830.

(15) (a) Stürzebecher, J.; Markwardt, F.; Voigt, B.; Wagner, G.; Walsmann, P. Cyclic amides of Nα-arylsulfonylamino acylated 4-amidinophenylalanine − tight binding of thrombin. *Thromb. Res*. **1983**, *29*, 635−642. (b) Stürzebecher, J.; Markwardt, F.; Voigt, B.; Wagner, G. Inhibition of bovine and human thrombins by derivatives of benzamidines. *Thromb. Res*. **1984**, *36*, 457−467.

(16) Groutas, W. C.; Kuang, R.; Venkataraman, R.; Epp, J. B.; Ruan, S.; Prakash, O. Structure-based design of mechanism-based inhibitors of the serine proteinases employing a novel amino acid-derived heterocyclic scaffold. *Biochemistry* **1997**, *36*, 4739−4750.

(17) (a) Supuran, C. T.; Briganti, F.; Scozzafava, A. Sulfenamido-sulfonamides as inhibitors of carbonic anhydrase isozymes I, II and IV. *J. Enzyme Inhib*. **1997**, *12*, 175−190. (b) Briganti, F.; Pierattelli, R.; Scozzafava, A.; Supuran, C. T. Carbonic anhydrase inhibitors Part 37. Novel classes of isozyme I and II inhibitors and their mechanism of action. Kinetic and spectroscopic investigations on native and cobalt-substituted enzymes. *Eur. J. Med. Chem*. **1996**, *31*, 1001−1010.

(18) (a) Supuran, C. T.; Ilies, M. A.; Scozzafava, A. Carbonic anhydrase inhibitors Part 29. Interaction of isozymes I, II and IV with benzolamide-like derivatives. *Eur. J. Med. Chem*. **1998**, *33*, 739−752. (b) Mincione, F.; Menabuoni, L.; Briganti, F.; Mincione, G.; Scozzafava, A.; Supuran, C. T. Carbonic anhydrase inhibitors: Inhibition of isozymes I, II and IV with *N*-hydroxysulfonamides − A novel class of intraocular pressure lowering agents. *J. Enyme Inhib*. **1998**, *13*, 267−284.

(19) Winnek, P. S.; Anderson, G. W.; Marson, H. W.; Faith, E.; Roblin, R. O., Jr. Studies in chemotherapy. V. Sulfanilylcyanamide and related compounds. *J. Am. Chem. Soc*. **1942**, *64*, 1682−1685.

(20) Supuran, C. T.; Scozzafava, A.; Briganti, F. Carbonic anhydrase inhibitors: *N*-Cyanosulfonamides, a new class of high affinity isozyme II and IV inhibitors. *J. Enyme Inhib.* **1999**, *14*, 289−306.
(21) Scozzafava, A.; Supuran, C. T. Carbonic anhydrase inhibitors. Arylsulfonylureido and arylureido-substituted aromatic and heterocyclic sulfonamides: towards selective inhibitors of carbonic anhydrase isozyme I. *J. Enzyme Inhib.* **1999**, *15*, 343−363.
(22) (a) Anderson, G. W.; Zimmerman, J. E.; Callahan, F. M. N-Hydroxysuccinimide esters in peptide synthesis. *J. Am. Chem. Soc.* **1963**, *85*, 3039. (b) Sheehan, J. C.; Ledis, S. L. Total synthesis of a monocyclic peptide lactone antibiotic, etamycin. *J. Am. Chem. Soc.* **1973**, *95*, 875−879. (c) Zlatoidsky, P.; Maliar, T. Synthesis of 4-(4-guanidinobenzoyloxy)benzamides and 1-(4-guanidinobenzoyloxy)benzoyloxy acetamides as trypsin inhibitors. *Eur. J. Med. Chem.* **1996**, *31*, 895−900.
(23) Lottenberg, R.; Christensen, U.; Jackson, C. M.; Coleman, P. L. Assay of coagulation proteases using peptide chromogenic and fluorogenic substrates. *Methods Enzymol.* **1981**, *80*, 341−361.
(24) Clare, B. W.; Scozzafava, A.; Briganti, F.; Iorga, B.; Supuran, C. T. Protease inhibitors: Synthesis and QSAR study of nonbasic thrombin inhibitors incorporating sulfonylguanidine moieties as S1 anchoring groups. *J. Enzyme Inhib.*, in press.
(25) Hyperchem v. 5.1; available from Hypercube, Inc., 1115 NW 4th St., Gainesville, FL 32601-4256.
(26) Stewart, J. J. P. MOPAC: A general molecular orbital package, version 6.0. *QCPE Bull.* **1990**, *10*, 86−87.
(27) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. AM1: A new general purpose quantum mechanical model. *J. Am. Chem. Soc.* **1985** *107*, 3902−3909.
(28) Stewart, J. J. P. MOPAC 93.00; Fujitsu Ltd.: Tokyo, Japan, 1993. Also Stewart, J. J. P. MOPAC93 Release 2. *QCPE Bull.* **1995,** *15*, 13−14 (copyright Fujitsu 1993, all rights reserved).
(29) Klamt, A.; Sshüürman, G. COSMO: A new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc., Perkin Trans. 2* **1993**, 799−805.
(30) Clare, B. W. The Frontier Orbital Phase Angles: Novel QSAR Descriptors for Benzene Derivatives, Applied to Phenylalkylamine Hallucinogens. *J. Med. Chem.* **1998**, *41*, 3845−386.
(31) Fukui, K.; Yonezawa, T.; Nagata, C. Theory of Substitution in Conjugated Molecules. *Bull. Chem. Soc. Jpn.* **1954**, *27*, 423−427.
(32) Kikuchi, O. Systematic QSAR Procedures with Quantum Chemical Descriptors *Quant. Struct.-Act. Relat.* **1987**, *6*, 179−184.
(33) ClogP for Windows, v. 1.0.0; available from BioByte Corp., 201 West 4th St., Suite 204, Claremont, CA 91711.
(34) BMDP Dynamic, release 7; available from BMDP Statistical Software, Inc., 1440 Sepulveda Blvd., Suite 316, Los Angeles, CA 90025.
(35) Furnival, G. M.; Wilson R. W. Regression by leaps and bounds. *Technometrics* **1974**, *16*, 499−511.
(36) Chatterjee, S.; Price, B. *Regression Analysis by Example,* 1st ed.; Wiley: New York, 1977; pp 199−200.
(37) Topliss, J. G.; Costello, R. J. Chance correlations in structure activity studies using multiple regression analysis. *J. Med. Chem.* **1972**, *15*, 1066−1068.
(38) Topliss, J. G.; Edwards, R. J. Chance Factors in Studies of Quantitative Structure−Activity Relationships. *J. Med. Chem.* **1979**, *22*, 1238−1244.
(39) Efroymson, M. A. Multiple Regression Analysis. In *Mathematical Methods for Digital Computers;* Ralston, A., Wilf, H. S., Eds.; Wiley: New York, 1962; p 191.
(40) Clare, B. W. Alternating conditional expectations in QSAR. In *Methods and Principles in Medicinal Chemistry*; Mannhold, R., Krogsgaard-Larsen, P., Timmerman, H., Eds.; VCH: Weinheim, 1994. *Advanced Computer-Assisted Techniques in Drug Discovery*; van de Waterbeemd, H., Ed.; pp 284−289.
(41) Clare, B. W. Structure−activity correlations for psychotomimetics. 2. Phenylalkylamines: a treatment of nonlinearity using the alternating conditional expectation technique *Chemom. Intell. Lab. Syst.* **1993**, *18*, 71−93.
(42) Banner, D. W.; Hadvary, P. Crystallographic analysis at 3.0 Å resolution of the binding to human thrombin of four active site-directed inhibitors. *J. Biol. Chem.* **1991**, *266*, 20085−20093.
(43) Gaussian 94, revision A.1; M. J. Frisch, G. W. Trucks, H. B. Schlegel, P. M. W. Gill, B. G. Johnson, M. A. Robb, J. R. Cheeseman, T. A. Keith, G. A. Petersson, J. A. Montgomery, K. Raghavachari, M. A. Al-Laham, V. G. Zakrzewski, J. V. Ortiz, J. B. Foresman, C. Y. Peng, P. Y. Ayala, M. W. Wong, J. L. Andres, E. S. Replogle, R. Gomperts, R. L. Martin, D. J. Fox, J. S. Binkley, D. J. Defrees, J. Baker, J. P. Stewart, M. Head-Gordon, C. Gonzalez, J. A. Pople; Gaussian, Inc.: Pittsburgh, PA, 1995.
(44) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37*, 785−789.
(45) Parr, R. G.; Yang, W. *Density-functional theory of atoms and molecules;* Oxford University Press: Oxford, 1989.
(46) Foresman, J. B.; Keith, T. A.; Wiberg, K. B.; Snoonian, J.; Frisch, M. J. Solvation Effects 5. Influence of Cavity Shape, Truncation of Electrostatics, and Electron Correlation on ab Initio Reaction Field Calculations. *J. Phys. Chem.* **1996**, *100*, 16098−16104.
(47) Pallas for Windows, v. 3.1 pK_a Prediction Module, p*K*alc 3.1; available from CompuDrug Chemistry Ltd., H-1395 Budapest, 62 P.O.B. 405, Hungary.
(48) Balasubramanian, N.; Laurent, D. R. S.; Federici, M. E.; Meanwell, N. A.; Wright, J. J.; Schumacher, W. A.; Seiler, S. M. Active site-directed synthetic thrombin inhibitors: synthesis, in vitro and in vivo activity profile of BYM 44621 and analogues. An examination of the role of the amino group in the D-Phe-Pro-Arg-H series. *J. Med. Chem.* **1993**, *36*, 300−303.
(49) Bode, W.; Turk, D.; Stürzebecher, J. Geometry of binding of benzamidine and arginine-based inhibitors NAPAP and MQPA to human α-thrombin: X-ray crystallographic determination of the NAPAP-trypsin complex and modeling of NAPAP-thrombin and MQPA-thrombin. *Eur. J. Biochem.* **1990**, *193*, 175−182.
(50) Czapinska, H.; Otlewski, J. Structural and energetic determinants of the S1-site specificity in serine proteases. *Eur. J. Biochem.* **1999**, *260*, 571−595.
(51) Stone, S. R.; Le Bonniec, B. F. Thrombin. In *Handbook of Proteolytic Enzymes, CD-ROM*; Barrett, A. J., Rawlings, N. D., Woessner, J. F., Eds.; Academic Press: London, 1998; Chapter 55.
(52) Bode, W.; Turk, D.; Karshikov, A. The refined 1.9-A X-ray crystal structure of D-Phe-Pro-Arg chloromethyl ketone-inhibited human α-thrombin: structure analysis, overall structure, electrostatic properties, detailed active-site geometry, and structure−function relationship. *Protein Sci.* **1992**, *1*, 426−471.
(53) Rezaie, A. R.; Esmon, C. T. Molecular basis of residue 192 participation in determination of coagulation protease specificity. *Eur. J. Biochem.* **1996**, *242*, 477−484.
(54) Clare, B. W. The frontier orbital phase angles: A theoretical interpretation. *J. Mol. Struct. (THEOCHEM),* in press.
(55) Gomez-Jeria, J. S. Calculation of the nucleophilic superdelocalizability by the CNDO method. *J. Pharm. Sci.* **1982**, *71*, 1423−1424.
(56) Bell, P. H.; Roblin, R. O. Studies in chemotherapy. VII. A theory of the relation of structure to activity of sulfanilamide type compounds. *J. Am. Chem. Soc.* **1942,** *64*, 2905−2917.
(57) Hemker, H. C.; Willems, G. M.; Beguin, S. A computer assisted method to obtain the prothrombin activation velocity in whole plasma independent of thrombin decay process. *Thromb. Haemostasis* **1986**, *56*, 9−17.

JM9903693