(7) E. L. Tolman, H. M. Tepperman, and J. Tepperman, *Am. J. Physiol.*, **218**, 1313 (1970).
(8) M. S. Schotz, A. Scanu, and I. H. Page, *Am. J. Physiol.*, **188**, 399 (1957).
(9) E. L. Bierman and D. Porte, Jr., *Ann. Intern. Med.*, **68**, 926 (1968).
(10) J. E. M. Groener and L. M. G. van Golde, *Biochim. Biophys. Acta*, **487**, 105 (1977).
(11) D. T. Witiak, E. Kuwano, D. R. Feller, J. R. Baldwin, H. A. I. Newman, and S. K. Sankarapa, *J. Med. Chem.*, **19**, 1214 (1976).
(12) D. T. Witiak, H. A. I. Newman, G. K. Poochikian, S. W. Fogt, J. R. Baldwin, C. L. Sober, and D. R. Feller, *J. Med. Chem.*, **21**, 833 (1978).
(13) R. C. Cavestri and D. T. Witiak, presented to the Division of Medicinal Chemistry, 8th Central Regional Meeting of the American Chemical Society, Akron, Ohio, May 21, 1976.
(14) J. R. Dyer in "Application of Absorption Spectroscopy of Organic Compounds", Prentice-Hall Foundations of Modern Organic Chemistry Series, K. L. Rinehart, Jr., Ed., Prentice-Hall, Englewood Cliffs, N.J., 1965, p 34.
(15) N. M. Yoom, C. S. Pak, H. C. Brown, S. K. Krishnamurthy, and T. P. Stocky, *J. Org. Chem.*, **38**, 2786 (1973).
(16) E. W. Collington and A. I. Meyers, *J. Org. Chem.*, **36**, 3044 (1971).
(17) D. T. Witiak, W. P. Heilman, S. K. Sankarappa, R. C. Cavestri, and H. A. I. Newman, *J. Med. Chem.*, **18**, 934 (1975).
(18) H. A. I. Newman, D. R. Feller, and D. T. Witiak, unpublished results.
(19) L. W. White, *Biochem. Pharmacol.*, **178**, 361 (1971).
(20) P. Segal, R. S. Roheim, and H. A. Eder, *J. Clin. Invest.*, **51**, 1632 (1972).
(21) K. W. Walton, P. J. Scott, J. V. Jones, R. F. Fletcher, and T. Whitehead, *J. Atheroscler. Res.*, **3**, 396 (1963).
(22) P. J. Scott and P. J. Hurley, *J. Atheroscler. Res.*, **9**, 25 (1969).
(23) N. J. Lewis, D. R. Feller, G. K. Poochikian, and D. T. Witiak, *J. Med. Chem.*, **17**, 41 (1974).
(24) D. L. Azarnoff, J. Reddy, C. Hignite, and T. Fitzgerald, *Proc. Int. Congr. Pharmacol., 6th, 1975*, 4, 137–147 (1976).
(25) L. F. Fieser and M. Fieser, "Reagents for Organic Synthesis", Vol. 1, Wiley, New York, N.Y., 1968, p 255.
(26) J. S. McIntyre, *Can. J. Chem.*, **45**, 767 (1967).
(27) The silicic acid used for adsorption column chromatography was prepared by the method of V. Suchy (personal communication) from Bratislava, Czechoslovakia. One pound of commercial silicic acid (100 mesh) was suspended in 4 L of $H_2O$. After standing for 1 h the suspended fines were removed by decanting. The settled solids were resuspended in 4 L of $H_2O$ by stirring (in such a way as to avoid a vortex) and allowed to settle for 1 min. The suspended material in the upper one-third volume of water was siphoned into a second beaker. This was repeated (six to eight times) until the upper one-third volume was essentially free of solids, affording grade I. The contents of the second beaker were resuspended and allowed to settle for 2 min (vide supra), affording grade II. The method may be repeated four additional times. With settling times of 4, 8, 16, and 32 min, grades III–VI are obtained, respectively. Grades I–III are useful for crude separations, whereas grades IV and V are most suited for high-resolution chromatography. Grade VI is useful for thin-layer chromatography. All six grades were filtered, air-dried, and activated at 110–130 °C for 12–24 h. The silicic acid can be deactivated by empirically adding small percentages (up to 10% w/w) of $H_2O$.
(28) M. Eggstein, *Klin. Wochenschr.*, 44, 267 (1966).
(29) W. R. Holub and F. A. Galli, *Clin. Chem. (Winston-Salem, N.C.)*, **18**, 239 (1972).
(30) C. C. Allain, L. Poon, S. G. Chan, W. Richmond, and P. Fu, *Clin. Chem. (Winston-Salem, N.C.)*, **20**, 470 (1974).
(31) N. J. Lewis, D. T. Witiak, and D. R. Feller, *Proc. Soc. Exp. Biol. Med.*, **145**, 281 (1974).
(32) A. H. Phillips and R. G. Langdon, *J. Biol. Chem.*, **237**, 2652 (1963).
(33) T. Kinoshita and S. Horie, *J. Biochem. (Tokyo)*, **61**, 26 (1967).
(34) O. Lowry, N. Rosenbrough, A. Farr, and R. Randall, *J. Biol. Chem.*, **193**, 265 (1951).
(35) L. L. Abell, B. B. Levy, B. B. Brodie, and F. E. Kendall, *J. Biol. Chem.*, **195**, 357 (1952).
(36) A. C. Parekh and D. H. Jung, *Anal. Chem.*, **42**, 1423 (1970).
(37) F. G. Soloni, *Clin. Chem., (Winston-Salem, N.C.)*, **17**, 529 (1971).

# Inhibition of Four Human Serine Proteases by Substituted Benzamidines

Judith M. Andrews,* Daniel P. Roman, Jr., David H. Bing,

*Center for Blood Research, Boston, Massachusetts 02115*

and Michael Cory

*Wellcome Research Laboratory, Research Triangle Park, North Carolina 27709.* *Received May 11, 1978*

A series of substituted benzamidines has been examined for their inhibitory activity against the human serine proteases—trypsin, thrombin, plasmin, and C1s̄, a subunit of the first component of complement. The inhibition constants obtained for each enzyme were correlated with physical-chemical properties of the substituent group using the quantitative structure–activity relationship approach. This analysis indicated that plasmin and C1s̄ are very similar in their interactions with substituted benzamidines. The binding of benzamidines in both enzymes was affected by electron donation from the substituent and its hydrophobicity. Thrombin–benzamidine interaction was affected only by the hydrophobicity of the substituent. Trypsin displayed a complex interaction with substituted benzamidines, and interaction was dependent on molar refractivity and molecular weight. Certain substituents deviated significantly from the interactions predicted by the analysis. These compounds, the (*m*- and *p*-amidinophenyl)pyruvic acids, when analyzed by computer modeling, suggested that direct interaction between the substituent and the enzyme surface is important in assessing the effect of substituent groups on inhibitory activity.

A central feature of blood coagulation, complement activation, fibrinolysis, and digestion is the activation of serine proteases[1] which specifically hydrolyze protein substrates. Although the mechanism of proteolysis by these enzymes has been extensively investigated,[2-4] the structural basis for substrate-binding specificity is poorly understood. We have undertaken studies to delineate the chemical basis of the substrate-binding specificity of four human serine proteases—trypsin (E.C. 3.4.4.4), thrombin (E.C. 3.4.4.13), plasmin (E.C. 3.4.4.14), and the complement enzyme C1s̄.[5] Three of the four enzymes, trypsin, thrombin, and plasmin, bind to protein substrates at lysine

0022-2623/78/1821-1202$01.00/0 

and/or arginine residues.[6-8] All four enzymes hydrolyze the same basic and aromatic amino acid esters[1,9-11] and are competitively inhibited by benzamidine, a small organic molecule which is an excellent model for the cationic side chains of arginine and lysine.[12-14] Despite their similarities, the four enzymes exhibit a wide range of physiologic function and protein-substrate specificity.

Human trypsin is a pancreatic, digestive enzyme consisting of a single polypeptide chain of 23 000 daltons.[15] It has a specificity toward lysyl and arginyl residues and cleaves the C-terminal to these residues in any protein. Human plasmin is a serum protease of the fibrinolytic system consisting of two polypeptide chains of 48 000 and 25 700 daltons.[8] Plasmin cleaves C-terminal to both lysine and arginine in a wide variety of proteins including casein, IgG, insulin, C1s̄, fibrin, and others.[8] Human thrombin, the proteolytic enzyme generated from the procoagulant factor prothrombin, contains two polypeptide chains of 32 000 and 4 500 daltons[16] and cleaves C-terminal to arginine in fibrinogen, its primary protein substrate.[17,18] C1s̄ is a subcomponent of the first component of complement and consists of two disulfide-linked polypeptide chains of 59 000 and 28 000 daltons.[19] Although the specific bond cleaved is not known, this enzyme displays very high protein-substrate specificity, cleaving only the complement proteins C2 and C4.[20]

In these studies we have determined $K_i$, the competitive-inhibition constant, of 21 substituted benzamidines for each of the four enzymes. $K_i$ provides a quantitative measure of the affinity of the enzyme for these structurally related molecules.[21] The inhibition data were analyzed by the quantitative structure–activity relationship (QSAR) approach. QSAR analysis was developed by Hansch[22] from the Hammett[23] equation which is used in organic chemistry to correlate chemical activity with electronic effects of substituents. In 1968, Hansch[24] expanded this approach to include hydrophobic interactions important in biological systems. Since that time many parameters have been used to characterize a wide variety of chemical interactions in biological systems. These phenomena ranged from tadpole narcosis[25] to inhibition of whole guinea pig complement[23,27] to substrate specificity of purified enzymes.[28,29] The relationships obtained from QSAR analysis provide a means to compare the structural basis for the different protein-substrate specificities exhibited by trypsin, thrombin, plasmin, and C1s̄.

## Experimental Section

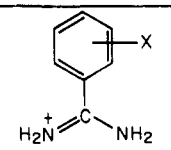
**Preparation of Enzymes.** Activated C1s̄ was isolated by affinity chromatography on IgG-(*p*-azobenzyloxyethylsulfonoethyl)-Sepharose 6B[30] by a modification of the method of Assimeh et al.[31] from Cohn fractions I + III (generously provided by Dr. Lewis Larson, Biological Laboratories, Jamaica Plain, MA). The fraction, eluted with 100 mM EDTA, pH 7.4, was used as the source of C1s̄ activity. It fulfilled the functional criteria for C1s̄, namely, catalysis of EAC4$\bar{2}$ formation in the sheep-erythrocyte cytolytic assay and induction of vascular leakage in guinea pig skin.[32] C1s̄ was stored in 100 mM EDTA (pH 7.4, 0 °C) at a concentration of 5–10 mg/mL, and a single preparation was used to obtain all of the inhibition constants reported.

Human plasmin and thrombin were the generous gift of Dr. Robert Rosenberg, Sidney Farber Cancer Center, Boston, MA. Plasminogen was prepared by the method of Deutsch and Mertz[33] and activated as described by Castellino et al.[34] Thrombin was purified by the method of Rosenberg and Waugh.[35]

Human trypsin purified from pancreas[15] was generously provided by Dr. James Travis, University of Georgia, Athens, GA.

**Preparation of Inhibitors.** All prepared compounds had melting points in agreement with the literature values or an elemental analysis (for C, H, and N) that was within 0.4% of the calculated value. Infrared spectra (Perkin-Elmer 137) were consistent with the assigned structures. The compounds tested are listed in Table I. Benzamidine hydrochloride hydrate (**7**), *m*-nitrobenzamidine hydrochloride (**4**), and *m*-aminobenzamidine hydrochloride (**8**) were purchased from Aldrich Chemical Co.; (*m*- and *p*-amidinophenyl)pyruvic acids (**19** and **15**) were prepared by the method of Richter et al.[36] The *p*-toluenesulfonates of *p*-methoxybenzamidine (**18**) and *p*-ethoxybenzamidine (**17**) were prepared from *p*-hydroxybenzamidine and the appropriate alkyl bromide by the method of Baker and Erickson.[37] The *m*- and *p*-(hydroxymethyl)benzamidines (**2** and **21**) were prepared by $NaBH_4$ reduction of *m*- and *p*-cyanobenzaldehyde to the respective benzyl alcohols which were converted to the amidines by the Pinner reaction.[38] 3-(Dimethylamino)benzamidine dipicrate (**10**) was prepared from 3-(dimethylamino)benzonitrile obtained by reductive alkylation of 3-aminobenzonitrile with $NaBH_3CN$.[39] The remaining compounds have been reported by Baker and Cory.[40,41]

Table I. Substituent Parameters Used to Develop the Equations in Table V[a]



| no. | X | $\pi$ | $\pi^2$ | $MR_m$[b] | $M_r$[b] | $R$ |
|---|---|---|---|---|---|---|
| 1 | 4-$NO_2$ | −0.28 | 0.08 | 0.10 | 4.5 | 0.16 |
| 2 | 3-$CH_2OH$ | −1.03 | 1.06 | 0.72 | 3.0 | 0 |
| 3 | 2-Me | 0.56 | 0.31 | 0.10 | 1.4 | −0.13 |
| 4 | 3-$NO_2$ | −0.28 | 0.08 | 0.74 | 4.5 | 0.16 |
| 5 | 3-$CO_2H$ | −0.32 | 0.10 | 0.69 | 4.4 | 0.15 |
| 6 | 3-$CH_2C_6H_5$ | 2.01 | 4.04 | 3.00 | 9.0 | −0.01 |
| 7 | H | 0.00 | 0.00 | 0.10 | 0 | 0 |
| 8 | 3-$NH_2$ | −1.23 | 1.51 | 0.54 | 1.5 | −0.68 |
| 9 | 3-$C_6H_5$ | 1.96 | 3.84 | 2.53 | 7.6 | −0.08 |
| 10 | 3-$N(CH_3)_2$ | 0.18 | 0.03 | 1.55 | 4.3 | −0.92 |
| 11 | 3-OMe | −0.02 | 0.00 | 0.79 | 3.0 | −0.51 |
| 12 | 3,4-$Me_2$ | 1.12 | 1.25 | 0.56 | 2.8 | −0.26 |
| 13 | 3-Br | 0.86 | 0.74 | 0.88 | 7.9 | −0.17 |
| 14 | 3,5-$Me_2$ | 1.12 | 1.25 | 1.13 | 2.8 | −0.26 |
| 15 | 4-$CH_2COCO_2H$ | −0.09 | 0.01 | 0.10 | 8.6 | −0.15 |
| 16 | 3-$O(CH_2)_3OC_6H_5$ | 2.33 | 4.97 | 4.37 | 15 | −0.57 |
| 17 | 4-OEt | 0.38 | 0.14 | 0.10 | 4.4 | −0.44 |
| 18 | 4-OMe | −0.02 | 0.00 | 0.10 | 3.0 | −0.51 |
| 19 | 3-$CH_2COCO_2H$ | −0.09 | 0.01 | 1.69 | 8.6 | −0.15 |
| 20 | 3-naphthamidine | 1.32 | 1.74 | 0.87 | 5.0 | 0.01 |
| 21 | 4-$CH_2OH$ | −1.03 | 1.06 | 0.10 | 3.0 | 0 |

[a] The values of the substituent constants used in the QSAR equations were taken from a compilation by Hansch.[45] The substituent parameters are defined as follows: $\pi$, hydrophobicity; $MR_m$, molar refractivity of a meta substituent; $M_r$, molecular weight; and $R$, the electronic resonance parameter. [b] $MR_m$ and $M_r$ have been multiplied by 0.1 to make them equiscalar with the other values.

**Assay Procedures.** The esterolytic activity of all four enzymes, alone or in the presence of inhibitors, was determined using one of two synthetic substrates, *p*-nitrophenyl *N*-carbobenzyloxy-L-tyrosinate (*N*-Z-L-Tyr-ONp) or *p*-nitrophenyl $N^{\alpha}$-carbobenzyloxy-L-lysinate ($N^{\alpha}$-Z-L-Lys-ONp). *N*-Z-L-Tyr-ONp was purchased from Nutritional Biochemical Corp. and used without further purification. $N^{\alpha}$-Z-L-Lys-ONp was prepared by acid hydrolysis of *p*-nitrophenyl $N^{\alpha}$-benzyloxycarbonyl-$N^{\epsilon}$-*tert*-butoxycarbonyl-L-lysinate (Sigma Chemical Co.).

The production of *p*-nitrophenol from these substrates was measured on a Gilford 240N spectrophotometer and recorded continuously on a Heath SR225B recorder. The cell compartment was maintained at 25 °C. The enzyme concentrations used in these assays were 1.0–5.0 μM. Hydrolysis of *N*-Z-L-Tyr-ONp with all four enzymes was determined by the method described by Bing,[10] and the production of *p*-nitrophenol was monitored at 410 nm assuming $E_M^{410} = 16\,600$. Hydrolysis of $N^{\alpha}$-Z-L-Lys-ONp was carried out in 50 mM $CH_3CO_2Na$ and 50 mM NaCl at pH 6.0 as the spontaneous hydrolysis of this substrate at pH 8.1 masked the enzymatic hydrolysis. The production of *p*-nitrophenol from $N^{\alpha}$-Z-L-Lys-ONp was monitored at 340 nm assuming $E_M^{340} = 5800$. A stock solution of 1 mM $N^{\alpha}$-Z-L-Lys-ONp was prepared in 50%

Table II. Kinetic Parameters for the Reaction of Human Trypsin, Thrombin, Plasmin, and C1s̄ with *N*-Z-L-Tyr-ONp and $N^{\alpha}$-Z-L-Lys-ONp[a]

| | $K_{m(app)}$, μM | $V_{max}$, μM/min | $k_{cat}$, $s^{-1}$ |
|---|---|---|---|
| | *N*-Z-L-Tyr-ONp | | |
| trypsin | 9.3 | 9.7 | 37 (a) |
| thrombin | 14 | 3.9 | 12 (b) |
| plasmin | 73 | 8.5 | 3.2 (b) |
| C1s̄ | 38 | 7.2 | 3.8 (a) |
| | $N^{\alpha}$-Z-L-Lys-ONp | | |
| trypsin | 17 | 9.6 | 96 (a) |
| thrombin | 13 | 6.0 | 5.3 (b) |
| plasmin | 26 | 4.7 | 10 (b) |
| C1s̄ | 4.2 | 12 | 0.6 (a) |

[a] The kinetic parameters $K_m$ and $V_{max}$ were determined from Lineweaver-Burk[42] plots to ±10%. $k_{cat}$ was determined using enzyme concentration based on (a) total Lowry protein[47] or (b) active-site titration with nitrophenyl guanidinobenzoate.[4]

methanol–$H_2O$ containing $10^{-2}$ M HCl and used within 4 h of preparation; 10–50 μL of this stock solution was used per milliliter of assay buffer. The concentration of each inhibitor was adjusted to give 10–80% inhibition. The competitive-inhibition constant $K_i$ [$-K_i = (1/V_{max} - y_{intercept})/slope$] was calculated by the method of Dixon[42] from previously determined values of $V_{max}$. Least-squares analysis of the inhibition data yielded correlation constants of 0.92 or greater.

**Analysis of Data.** The $K_i$ values were treated using a standard regression analysis of the form $\log 1/K_i = A + BX_1 + CX_2 + ...$ where $X_1$, $X_2$, ... are the substituent parameters listed in Table I. The analysis was carried out on the NIH Prophet computer with programs available through the Prophet system.[43,44] The analysis involves the stepwise addition or deletion of parameters based on an *F* test which shows significance at the 95% level. Most of the substituent constants ($\pi$, MR, *F*, *R*, $\sigma$, $M_r$) were taken from the recent compilation by Hansch.[45] In addition to these, $\chi$, the connectivity constant of Kier,[46] and $\pi_{NO_2}$ were also considered. The separation of the meta and para contributions of the substituents and the dummy steric parameter, *D*, were also dealt with in this analysis. The values of the substituent constants relevant to this analysis are listed in Table I.

**Computer Trypsin Model.** The interactions of compounds **15** and **19** with trypsin were investigated using a computer model of trypsin.[47] This work was done in conjunction with Dr. Richard Feldman of the NIH Molecular Graphics Section.[48] The trypsin model used is available through Amscom (Atlas of Macromolecular Structure on Microfiche, Tracor Jitco, Inc., Rockville, MD). Analysis was done on a PDP-10 computer and the inhibitor was inserted into the enzyme site with the aid of interactive computer graphics in real time. Care was taken that no van der Waals radii (2.0 Å) overlapped and a minimum energy model of the inhibitor was constructed with the NIH Prophet computer.[43,44]

## Results

**Enzyme Kinetics.** The rate of reaction of each of the four enzymes in this study was proportional to the enzyme concentration over a fivefold range with 30 μM *N*-Z-L-Tyr-ONp or $N^{\alpha}$-Z-L-Lys-ONp. At constant enzyme concentration, substrate inhibition was observed for thrombin and plasmin as the substrate concentration was increased from 10 to 50 μM. For thrombin, substrate inhibition was observed at high concentrations of *N*-Z-L-Tyr-ONp but not at comparable concentrations of $N^{\alpha}$-Z-L-Lys-ONp. Plasmin was inhibited by high concentrations of $N^{\alpha}$-Z-L-Lys-ONp but not by similar concentrations of *N*-Z-L-Tyr-ONp. A Lineweaver–Burk plot[42] was used to determine $V_{max}$, $K_{m(app)}$, and $k_{cat}$ for each enzyme (Table II).

**Inhibition Constants.** The competitive-inhibition constants obtained for each of the enzymes are listed in Table III. The inhibition constants for trypsin, plasmin, and C1s̄ were determined using *N*-Z-L-Tyr-ONp, while the constants for thrombin were obtained with $N^{\alpha}$-Z-L-Lys-ONp. For trypsin, thrombin, and C1s̄ the inhibition constants were determined for selected compounds with both substrates. The compounds were selected to cover the maximum range of $K_i$ values. The $K_i$ values obtained for hydrolysis of the lysine ester were found to be identical (within experimental error) with those obtained for the tyrosine ester (Table IV). This indicates that *N*-Z-L-Tyr-ONp and $N^{\alpha}$-Z-L-Lys-ONp are binding at the same site in these enzymes regardless of the presence or absence of a cationic group on the substrate.

Table III. Observed and Calculated Values for Equations in Table V[a]

| | | log $1/K_i$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | trypsin | | thrombin | | plasmin | | C1s̄ | |
| | group | obsd | calcd | obsd | calcd | obsd | calcd | obsd | calcd |
| 1 | 4-$NO_2$ | 3.9 | 3.8 | 2.5 | 2.7 | 2.4[c] | 3.0 | 1.0[c] | 2.7 |
| 2 | 3-$CH_2OH$ | 4.5 | 4.7 | 2.6 | 2.4 | 2.6 | 3.0 | 2.0 | 2.6 |
| 3 | 2-Me | 3.8 | 4.5 | 1.0[c] | 3.0 | 3.5 | 3.5 | 2.8 | 3.4 |
| 4 | 3-$NO_2$ | 4.1 | 4.3 | 2.6 | 2.7 | 3.0 | 3.0 | 2.8 | 2.7 |
| 5 | 3-$CO_2H$ | 2.3[c] | 4.3 | 2.7 | 2.6 | 3.4 | 3.0 | 2.2 | 2.7 |
| 6 | 3-$CH_2C_6H_5$ | 5.2 | 5.3 | 3.4 | 3.6 | 3.9 | 3.7 | 3.2 | 3.8 |
| 7 | H | 5.1 | 4.9 | 2.9 | 2.8 | 3.3 | 3.2 | 3.2 | 3.0 |
| 8 | 3-$NH_2$ | 5.2 | 4.9 | 4.4[c] | 2.3 | 4.0 | 3.7 | 3.2 | 3.2 |
| 9 | 3-$C_6H_5$ | 5.4 | 5.2 | 3.7 | 3.6 | 3.5 | 3.8 | 3.6 | 3.9 |
| 10 | 3-$N(CH_3)_2$ | N.T.[d] | 5.1 | 3.1 | 2.9 | 4.2 | 4.3 | 3.7 | 4.1 |
| 11 | 3-OMe | 4.9 | 4.8 | 3.1 | 2.8 | 3.5 | 3.8 | 3.8 | 3.6 |
| 12 | 3,4-$Me_2$ | 5.1 | 4.6 | 2.8 | 3.2 | 2.5[c] | 3.8 | 3.9 | 3.8 |
| 13 | 3-Br | N.T. | 3.7 | 2.8 | 3.1 | 2.1[c] | 3.6 | 4.0 | 3.5 |
| 14 | 3,5-$Me_2$ | 4.8 | 5.1 | 2.0[c] | 3.2 | 3.6 | 3.8 | 4.1 | 3.8 |
| 15 | 4-$CH_2COCO_2H$ | 5.0 | 2.8 | 5.4[c] | 2.8 | 4.9[c] | 3.4 | 4.9[c] | 3.1 |
| 16 | 3-$O(CH_2)_3OC_6H_5$ | 5.0 | 5.0 | 3.8 | 3.7 | 4.7 | 4.4 | 5.0 | 4.6 |
| 17 | 4-OEt | 4.0 | 3.8 | 2.8 | 3.0 | N.T. | 3.8 | <1[b] | 3.6 |
| 18 | 4-$OCH_3$ | 4.0 | 4.2 | 2.7 | 2.8 | N.T. | 3.8 | <1[b] | 3.6 |
| 19 | 3-$CH_2COCO_2H$ | 3.1 | 4.2 | <1[b] | 2.8 | 2.8 | 3.3 | <1[b] | 3.1 |
| 20 | 3-naphthamidine | N.T. | 4.3 | 4.0 | 3.3 | 4.8[c] | 3.5 | <1[b] | 3.1 |
| 21 | 4-$CH_2OH$ | 4.1 | 4.2 | 2.5 | 2.4 | 3.2 | 3.0 | N.T. | 2.6 |

[a] The observed $K_i$ values were determined by the method of Dixon[42] using the $V_{max}$ values in Table II. [b] A $\log 1/K_i$ value of <1.0 indicates a compound for which the limit of solubility prevented an accurate determination of a $K_i$ over 100 mM. However, these compounds do show some inhibitory activity toward the enzyme. These compounds were not used to generate the equations on Table V. [c] These compounds appear to have a unique interaction with the respective enzymes and were eliminated from the equations in Table V. [d] N.T. = not tested.

Table IV. Comparison of Log $1/K_i$ Obtained Using *N*-Z-L-Tyr-ONp and $N^{\alpha}$-Z-L-Lys-ONp[a]

| R | log $1/K_i$ *N*-Z-L-Tyr-ONp | log $1/K_i$ $N^{\alpha}$-Z-L-Lys-ONp |
|---|---|---|
| trypsin | | |
| H | 5.1 | 5.2 |
| 3-$O(CH_2)_3OC_6H_5$ | 5.9 | 5.3 |
| 3-$NH_2$ | 5.2 | 5.1 |
| 3-$NO_2$ | 4.1 | 4.1 |
| C1$\bar{s}$ | | |
| H | 3.2 | 3.3 |
| 3-$O(CH_2)_3OC_6H_5$ | 4.8 | 4.5 |
| 3-$NH_2$ | 3.3 | 3.6 |
| 3-$NO_2$ | 2.7 | 2.6 |
| thrombin | | |
| H | 3.0 | 2.8 |
| 3-$O(CH_2)_3OC_6H_5$ | 4.1 | 3.8 |
| 3,4-$Me_2$ | 3.0 | 2.8 |
| 3-$C_6H_5$ | 3.8 | 3.7 |
| 3,5-$Me_2$ | 2.0 | 2.0 |

[a] A comparison was made of the inhibition constants obtained with *N*-Z-L-Tyr-ONp and $N^{\alpha}$-Z-L-Lys-ONp. Log $1/K_i$ (±0.1) values were obtained with each substrate for several substituted benzamidines with trypsin, C1$\bar{s}$, and thrombin.

**QSAR Analysis.** The observed values of log $1/K_i$ presented in Table III were correlated by multiparameter regression analysis with the substituent constants listed in Table I, and the correlations obtained are shown in Table V. The equations in Table V describe the chemical nature of the interaction between the benzamidines and each enzyme. The interactions of plasmin and C1$\bar{s}$ with benzamidines were almost identical as delineated by QSAR analysis. The inhibitory capacity of the benzamidines increased for both enzymes as resonance donation ($R$) of electron density into the phenyl ring and the hydrophobicity of the substituent increased. The binding of substituted benzamidines to thrombin was dependent only on the hydrophobicity of the substituent, the more hydrophobic the substituent the better the binding. Trypsin–benzamidine binding was enhanced by substituents with high molar refractivity and low molecular weight.

In the initial development of the QSAR equations all of the compounds with measurable $K_i$ values were considered. Those compounds where $K_i$ was too high to be determined (see footnote *b* in Table III) were omitted from the initial analysis. Log $1/K_i$ values calculated from these equations revealed a small number of compounds in which the observed values differed by 1 log or more from the calculated values. A second regression analysis was carried out in which these compounds were deleted, one at a time, from the equation. Those compounds (see footnote *c* in Table III) eliminated from the equations in Table V were determined by maximizing both $n$, the number of compounds, and $r^2$, the square of the correlation coefficient. In this manner those inhibitors which displayed unique behavior were identified. This does not alter the substituent parameters of the initial equation but results in small changes in the multiplying coefficients.

## Discussion

Benzamidine and a number of its derivatives are good inhibitors of the esterase activity of human trypsin, thrombin, plasmin, and C1$\bar{s}$. The primary force governing the binding of benzamidine to these enzymes is an ionic interaction between the positively charged amidine moiety and a negatively charged amino acid side chain in the enzyme. For trypsin and thrombin this amino acid side chain is the carboxyl group of an aspartate. The binding constant, $1/K_i$, observed for the substituted benzamidines varies widely for a given enzyme, depending on the nature of the substituent group. In most cases we attribute the variation in $1/K_i$ to the effect of the substituent on the benzamidine interaction. However, with some substituents these changes in $1/K_i$ may be the result of direct interactions between the enzyme and the individual substituent at a site adjacent to the anionic binding site. A similar interpretation was made by Thompson and Blout in their studies on the interaction of alanine polypeptide inhibitors of elastase.[49] The use of inhibition data to obtain information on the relations between enzyme specificity and structure posed some difficulties. Attempts to find similarities among the enzymes by comparing the magnitude of the inhibition constants yielded little information. The correlations obtained among the enzymes ranged from a high of 0.72 for plasmin and thrombin to a low of 0.27 for trypsin and plasmin.

An alternative method for analyzing inhibition data is to compare the inhibitory properties of the various compounds for a single enzyme and determine the physical–chemical properties which make the compounds better inhibitors. The equations generated by QSAR analysis are an indication of the properties of the substituent groups which make a particular molecule a better or worse inhibitor. This analysis yielded certain substituent properties which appear to have a significant effect on the inhibitory capacity of the parent benzamidine. However, QSAR analysis treats the effect of the substituent on the whole molecule and does not reflect any direct interaction of the substituent with the enzyme surface. To summarize the results of Table V, the binding of benzamidines to plasmin and C1$\bar{s}$ was enhanced by substituents which donate electrons into the aromatic ring through resonance as indicated by the negative dependence on $R$ (Table V). Binding is also enhanced to a lesser extent by more hydrophobic substituents. The interaction of thrombin with benzamidines is enhanced by the presence of hydrophobic substituents. Trypsin has a more complex interaction with benzamidines than the serum enzymes. Binding is enhanced by increasing the molecular weight of the substituent. The step by step derivation of the trypsin equation is shown in Table VI.[43,44] The parameter of choice for a single-parameter equation is $\pi$,[2] which gives

Table V. QSAR Equations Generated for Trypsin, Thrombin, Plasmin, and C1$\bar{s}$[a]

| | | $n$ | $s$ | $r^2$ |
|---|---|---|---|---|
| trypsin | $\log 1/K_i = 4.77\ (\pm 0.16) + 0.877\ (\pm 0.33)\ MR_m - 0.239\ (\pm 0.11)\ M_r$ | 15 | 0.32 | 0.71 |
| thrombin | $\log 1/K_i = 2.80\ (\pm 0.16) + 0.039\ (\pm 0.14)\ \pi$ | 15 | 0.29 | 0.70 |
| plasmin | $\log 1/K_i = 3.23\ (\pm 0.16) + 0.25\ (\pm 0.12)\ \pi - 1.11\ (\pm 0.43)\ R$ | 14 | 0.26 | 0.79 |
| C1$\bar{s}$ | $\log 1/K_i = 2.99\ (\pm 0.29) + 0.41\ (\pm 0.22)\ \pi - 1.11\ (\pm 0.75)\ R$ | 14 | 0.45 | 0.67 |

[a] The QSAR equations were generated by a standard multiparameter regression analysis of the form $\log 1/K_i = A + BX_1 + CX_2 + \ldots$ using the number of compounds ($n$) indicated. Those compounds eliminated include inhibitors with a $K_i$ too large to determine accurately (footnote *b* in Table III) and those eliminated by maximizing $n$ and $r^2$, the square of the correlation coefficient (footnote c in Table III). The numbers in parentheses in the equations are the 95% confidence limits, $s$ is the standard error between the observed and calculated values for the enzyme, and $F$ is the $F$ test overall; $F_{14,1;\alpha=0.1} = 3.85$.

Table VI. Stepwise Development of the Trypsin Equation in Table V[a]

| intercept | $\pi^2$ | $M_r$ | $MR_m$ | $F$ of entering parameter | total $F$ | $r^2$ | $s$ |
|---|---|---|---|---|---|---|---|
| 4.34 | 0.21 | | | 7.5 | 7.5 | 0.37 | 0.46 |
| 4.57 | 0.44 | −0.13 | | 5.8 | 8.1 | 0.58 | 0.39 |
| 4.75 | 0.12 | −0.23 | 0.70 | 6.3 | 9.8 | 0.72 | 0.33 |
| 4.77 | | −0.24 | 0.88 | 0.64 | 14.9 | 0.71 | 0.32 |

[a] The stepwise development of the QSAR equation is illustrated here. Step 1, a single parameter $\pi^2$ best describes the observed log $1/K_i$ values. Step 2, the addition of the molecular weight ($M_r$) parameter significantly improves $r^2$ (the square of the correlation coefficient) and $s$ (the standard error). Step 3, addition of a third parameter, $MR_m$ (meta molar refractivity), gives further improvement in $r^2$ and $s$. Step 4, the original parameter $\pi^2$ can be eliminated from the equation with a very small decrease in $r^2$ and a significant increase in the total $F$ statistic.

Table VII. Correlation Matrix for Variables in Regression Equations[a]

| | $\pi$ | $R$ | $\pi^2$ | $MR_m$ | $M_r$ |
|---|---|---|---|---|---|
| $\pi$ | 1.00 | 0.003 | 0.504 | 0.504 | 0.333 |
| $R$ | | 1.00 | 0.003 | 0.040 | 0.002 |
| $\pi^2$ | | | 1.00 | 0.712 | 0.381 |
| $MR_m$ | | | | 1.00 | 0.619 |
| $M_r$ | | | | | 1.00 |

[a] The relationship of the various substituent parameters to each other is illustrated in this table. The correlations are reported as $r^2$, the square of the correlation coefficient. The values in Table I were used to determine $r^2$.

an $r^2$ of 0.37. As the other parameters $M_r$ and $MR_m$ are added to the equation, the correlation coefficient and the total $F$ statistic increase significantly. After the addition of $M_r$ and $MR_m$ to the equation, the $\pi^2$ term can be deleted without a significant decrease in $r_2$.

Many of the parameters used in this analysis vary in a collinear manner and reflect similar properties of the substituents. Table VII is a correlation of the various parameters. From this table it is evident that $\pi$, $MR_m$, and $M_r$ are very similar. $R$, on the other hand, is entirely independent of these three properties. The collinearity between $\pi$ and $MR_m$ and $M_r$ suggests that thrombin is more like trypsin than like either of the other two enzymes. Trypsin is an enzyme with very broad substrate specificity and this difference is reflected in the $1/K_i$ values which are 1–2 logs higher for trypsin than for the other enzymes. It is possible that the binding of these inhibitors to trypsin is more sensitive to variations in the benzamidine substituent and small changes in a $K_i$ are more readily detected.

The equations listed in Table V can be compared with those of a similar analysis reported by Coats[27] for human plasmin, bovine thrombin, trypsin, and guinea pig complement. For the 3-substituted benzamidines Coats found that hydrophobicity was the single most important parameter. Other parameters which were found to be important were polarizability, a parameter similar to $\pi$, and the Hammett $\sigma_m$ parameter, an electronic term. Coats found that complement and thrombin have a negative dependence on $\sigma_m$ and plasmin and trypsin have a positive dependence. However, substituents in Coats' analysis have almost identical values for the $\sigma_m$ parameter. In a similar analysis of the inhibition of plasmin, trypsin, and thrombin by substituted benzamidines, Stürzebecker et al.[29] also reported that hydrophobicity and resonance were important parameters. In the present work, the variation in the electronic parameter, $R$, is much greater than that in the $\sigma_m$ parameter used by Coats and places more credibility on the dependence found in the electronic contribution reported here. In addition, previous QSAR analysis of inhibition of the complement system was carried out on partially purified complement components or the entire complement system of 11 distinct proteins. The data presented here were obtained using highly purified enzyme preparations of a single component, C1s̄.

Coats attributed the increased inhibition with electron-donating groups to a stabilization of the positive charge on the amidine residue. The separation of the Hammett parameter into resonance and field effects indicates that the electron donation by the substituent is a resonance contribution. Classical resonance does not allow a donation by a meta substituent to the amidine moiety, and, thus, the cationic group of these benzamidines would be affected little, if at all, by resonance donation of a meta substituent. Hansch[50] has suggested that $R$ will be important only when the resonance does not affect the ionic interaction. Our hypothesis is that these substituents increase the electron density in the benzene ring and are enhancing a charge-transfer type of interaction (possibly with tryptophan-215)[47] of the inhibitor with the enzyme surface. This idea is supported by the work of East and Trowbridge[51] who have reported difference spectra indicating a perturbation of the $\pi$ system of benzamidine on binding to bovine trypsin.

The best evidence for the validity of the interactions described by these equations is their compatibility with known structures. It was originally proposed by Krieger et al.[52] that only para-substituted benzamidines would inhibit trypsin. It has recently been reported by Bode and Schwager[53] that meta-substituted benzamidines will indeed fit the trypsin-binding pocket. We have found that meta-substituted benzamidines are good inhibitors for all of the enzymes and that $MR_m$ is an important parameter in describing the trypsin–inhibitor interaction.

While the QSAR equations in Table V reflect the inhibitory effects of the majority of the substituents, a few substituents exhibit anomalous behavior with respect to these equations. One example of anomalous behavior is the interaction of para-substituted benzamidines with C1s̄. Three of the five para-substituted benzamidines, $p$-$NO_2$ (**1**), $p$-OEt (**17**), and $p$-OMe (**18**), have observed $1/K_i$ values 2 logs smaller than those calculated from eq 4 (Table V). 3,4-Dimethylbenzamidine (**12**) has a $1/K_i$ value essentially identical with the value calculated from the equation. One possible explanation of the low inhibitory activity of the three para-substituted benzamidines might be that the C1s̄ binding site is sterically blocked from binding para substituents larger than a methyl group. This steric effect is not observed for the other enzymes in which the inhibitory effect of both meta and para substituents is amenable to analysis by a single QSAR equation.

The (*m*- and *p*-amidinophenyl)pyruvic acids (**19** and **15**) represent another set of anomalous compounds. The $1/K_i$ values obtained for the para compound were 1.5–2.0 logs higher than the values calculated from the equations in Table V, while the $1/K_i$ values obtained for the meta compound were lower than the calculated values. The pyruvic acid substituents at first appear to resemble structurally the synthetic esters (*N*-Z-L-Tyr-ONp and $N^{\omega}$-Z-L-Lys-ONp) used in this study. However, it has been estimated that these compounds exist 90% in the enol form.[54] The anomalous behavior of these compounds suggests that there may be some direct interaction between the pyruvic acid substituent and the enzyme surface. This type of interaction is not accounted for in the QSAR

analysis. To investigate the direct substituent–enzyme interaction both compounds were inserted into a computer model of bovine trypsin as determined by Huber.[47] The computer fit indicated that when the amidine of (*p*-amidinophenyl)pyruvic aicd was bound to the aspartate in the trypsin binding site the pyruvic acid moiety fell directly on the catalytic center. The enol carbon was adjacent to the hydroxyl of the active-site serine while the carboxylic acid function was adjacent to histidine-57. In contrast, when the (*m*-amidinophenyl)pyruvic acid was inserted into the trypsin binding site, the pyruvic acid moiety was not in a position to interact with the enzyme surface. It is the unique stereochemistry of the (*p*-amidinophenyl)pyruvate which leads to its unusual reactivity.

As a final note, we agree that QSAR analysis is a useful tool in determining the physical–chemical interactions which govern the binding of a series of structurally related compounds. Although this technique is not capable of predicting anomalies, it is a valuable tool in detecting anomalies such as those observed for the (amidinophenyl)pyruvic acids. These anomalies are a reminder that enzymes are complex systems and it is often the exception, rather than the rule, which explains their unique behavior.

**Acknowledgment.** The authors gratefully acknowledge Joseph P. Falke for his technical assistance and Rachelle Rosenbaum for typing the manuscript. This work was supported by National Institutes of Health Grants AM 17351, HL 18825, CA 17376, AI 16392, AI 100264, and AI 14779. D.H.B. is a recipient of an Established Investigatorship of the American Heart Association.

## References and Notes

(1) K. A. Walsh, "Proteases and Biological Control", E. Reich, D. B. Rifkin, and E. Shaw, Eds., Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1975, p 1.
(2) R. M. Stroud, M. Kreiger, R. E. Keoppe III, A. A. Kossiakoff, and J. L. Chambers in ref 1, p 13.
(3) B. Keil, "The Enzymes", Vol. III, P. D. Boyer, Ed., Academic Press, New York, N.Y., 1971, p 250.
(4) T. Chase, Jr., and E. Shaw, *Methods Enzymol.*, **19**, 20 (1970).
(5) The terminology for complement proteins is that suggested in *Bull. W. H. O.*, **39**, 935 (1968).
(6) K. A. Walsh, *Methods Enzymol.*, **19**, 41 (1970).
(7) T. Takayi and R. F. Doolittle, *Biochemistry*, **13**, 750 (1974).
(8) K. C. Robbins and L. Summaria, *Methods Enzymol.*, **19**, 184 (1970).
(9) N. R. Cooper, "Progress in Immunology", C. B. Amos, Ed., Academic Press, New York, N.Y., 1970, p 568.
(10) D. H. Bing, *Biochemistry*, **8**, 4053 (1969).
(11) D. T. Elmore and E. F. Curragh, *Biochem. J.*, **86**, 9P (1963).
(12) M. Mares-Guia and E. Shaw, *J. Biol. Chem.*, **240**, 1579 (1965).
(13) B. R. Baker and E. H. Erickson, *J. Med. Chem.*, **10**, 1123 (1967).
(14) T. J. Ryan, J. W. Fenton, II, T. Chang, and R. D. Feinman, *Biochemistry*, **15**, 1337 (1976).
(15) D. Johnson and J. Travis, *Anal. Biochem.*, **72**, 573 (1976).
(16) K. Bailey and F. A. Bettelheim, *Biochim. Biophys. Acta*, **18**, 495 (1955).
(17) I. Cohen, Z. Bohak, A. deVries, and W. Katchalski, *Eur. J. Biochem.*, **10**, 388 (1969).
(18) V. Mutt, S. Magnusson, J. E. Jorpes, and E. Dahl, *Biochemistry*, 4, 2358 (1965).
(19) R. T. Ziccardi and N. R. Cooper, *J. Immunol.*, **116**, 504 (1976).
(20) H. J. Müller-Eberhard in ref 9, 1971, p 568.
(21) J. G. Morris, "A Biologist's Physical Chemistry", 2nd ed., Addison-Wesley, Reading, MA, 1974, p 296.
(22) C. Hansch, *Adv. Chem. Ser.*, **No. 114**, 20 (1970).
(23) L. P. Hammett, "Physical Organic Chemistry", 2nd ed., McGraw-Hill, New York, N.Y., 1970.
(24) C. Hansch, A. R. Steward, S. M. Anderson, and D. Bentley, *J. Med. Chem.*, **11**, 1 (1968).
(25) C. Hansch and M. Clayton, *J. Pharm. Sci.*, **62**, 1 (1973).
(26) C. Hansch and M. Yoshimoto, *J. Med. Chem.*, **17**, 1160 (1974).
(27) E. A. Coats, *J. Med. Chem.*, **16**, 1102 (1973).
(28) C. Hansch, C. Grieco, C. Silipo, and A. Vittoria, *J. Med. Chem.*, **20**, 1420 (1977).
(29) J. Stürzebecker, J. F. Markwardt, G. Wagner, and P. Walsmann, *Acta Biol. Med. Ger.*, **35**, 1665 (1976).
(30) J. M. Andrews, R. H. Painter, and D. H. Bing, unpublished results.
(31) S. Assimeh, D. H. Bing, and R. H. Painter, *J. Immunol.*, **112**, 225 (1974).
(32) J. A. Andrews, F. S. Rosen, S. J. Silverberg, M. Cory, E. E. Schneeberger, and D. H. Bing, *J. Immunol.*, **118**, 466 (1977).
(33) G. D. Deutsch and E. T. Mertz, *Science*, **170**, 1095 (1970).
(34) F. J. Castellino, J. M. Sodetz, W. J. Brockway, and G. E. Siefring, Jr., *Methods Enzymol.*, **45**, 244 (1976).
(35) R. D. Rosenberg and D. F. Waugh, *J. Biol. Chem.*, **245**, 5049 (1970).
(36) P. Richter, H. G. Kazmirowski, G. Wagner, and C. H. Garbe, *Pharmazie*, **28**, 585 (1973).
(37) B. R. Baker and E. H. Erickson, *J. Med. Chem.*, **12**, 112 (1969).
(38) A. W. Dox and F. C. Whitmore, "Organic Syntheses", Collect. Vol. I, H. Gilman and A. H. Blatt, Eds., Wiley, New York, N.Y., 1941, p 5.
(39) R. F. Borch, *Org. Synth.*, **52**, 124 (1972).
(40) B. R. Baker and M. Cory, *J. Med. Chem.*, **12**, 1053 (1969).
(41) B. R. Baker and M. Cory, *J. Med. Chem.*, **12**, 1049 (1969).
(42) M. Dixon and E. C. Webb, "Enzymes", 2nd ed., Academic Press, New York, N.Y., 1964, p 329.
(43) *Fed. Proc., Fed. Am. Soc. Exp. Biol.*, **32**, 1744 (1973).
(44) *Chem. Eng. News*, **51**, 20 (Aug 20, 1973).
(45) C. Hansch, A. Leo, S. H. Unger, K. H. Kim, D. Nikaitani, and E. J. Lien, *J. Med. Chem.*, **16**, 1207 (1973).
(46) L. B. Kier, L. H. Hall, W. J. Murray, and M. Randie, *J. Pharm. Sci.*, **64**, 1970 (1975).
(47) R. Huber, D. Kukla, W. Bode, P. Schwager, K. Bartels, J. Desenhofer, and W. Stergman, *J. Mol. Biol.*, **89**, 731 (1974).
(48) R. A. Feldmann, D. H. Bing, B. F. Furie, and B. Furie, submitted for publication in *Proc. Natl. Acad. Sci. U.S.A.*
(49) R. C. Thompson and E. R. Blout, *Proc. Natl. Acad. Sci. U.S.A.*, **67**, 1734 (1970).
(50) C. Hansch, *Acc. Chem. Res.*, **11**, 1 (1969).
(51) E. J. East and C. G. Trowbridge, *Arch. Biochem. Biophys.*, **125**, 334 (1968).
(52) M. Krieger, L. M. Kay, and R. M. Stroud, *J. Mol. Biol.*, **83**, 209 (1974).
(53) W. Bode and P. Schwager, *J. Mol. Biol.*, **98**, 693 (1975).
(54) R. Richter and C. Wagner, *Pharmazie*, **31**, 707 (1976).