*Tetrahedron* Vol. 47, No. 14/15, pp. 2519-2534, 1991
Printed in Great Britain

0040-4020/91 $3.00+.00


# LONG RANGE ELECTROSTATIC EFFECTS IN PEPSIN CATALYSIS

Petr Kuzmič, Chong-Qing Sun, Zhi-Cheng Zhao, and Daniel H. Rich *

School of Pharmacy and Department of Chemistry
University of Wisconsin-Madison
425 N. Charter St., Madison, WI 53706

(*Received in USA* 24 *September* 1990)

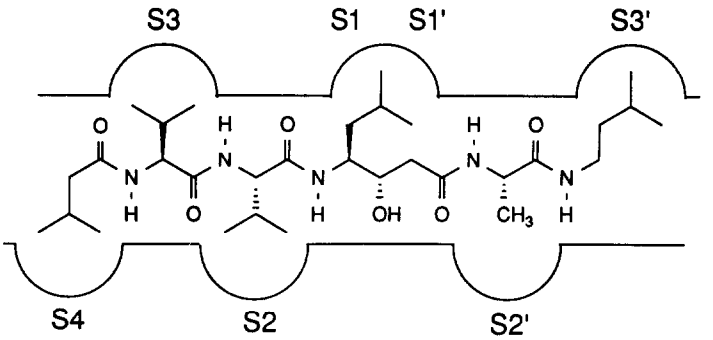
**Abstract:** Binding of polycationic substrates and inhibitors to pepsin is a stepwise process. Fast nonspecific association, governed by long range electrostatic forces, is followed by slow surface diffusion into the active site.

## Introduction

Proteinases have served a particularly important role in the emergence of basic ideas currently invoked to rationalize enzymatic catalysis. For example, the concept of secondary specificity[1] was derived in part from the kinetics of oligopeptide substrates of pepsin. In a series of classical studies, Fruton and coworkers[2] characterized the influence of various aminoacids in the substrate sequence on the steady state kinetic parameters. These results have shown that reactivity is usually increased by lengthening the peptide chain, and/or by increasing the hydrophobic nature of the aminoacids on either side of the cleaved peptide bond. These effects can be explained in terms of stabilization of the productive enzyme-substrate complex by hydrophobic contacts throughout the extended catalytic cleft. The preference of pepsin for hydrophobic ligands was also demonstrated on low molecular weight inhibitors, such as substituted benzene derivatives[3], aliphatic alcohols[4], and aliphatic acids[5]. For all three classes of compounds it has been shown that the logarithm of inhibition constant, and thus the free energy of binding, is linearly dependent on hydrophobic energy. A close look at some previously reported peptide inhibitors also indicates a dominant role of hydrophobic binding; the logarithm of the inhibition constant for compounds of the general formula $R_1R_2CHCO$-Val-Sta-Ala-Iaa[6] linearly correlates with the total number of carbon atoms in alkyl groups $R_1$ and $R_2$. This peptide was structurally varied in the position $P_3$ (notation of Schechter and Berger[7]), thus the linear correlation testifies for predominantly hydrophobic character of the binding subsite $S_3$ on the enzyme (Scheme 1).

It is a common practice to correlate structural and kinetic properties of enzymatic substrates with an implied image of the ligand *bound in the active site*. From this specific binding model[8], in conjunction with most of the previously reported data on pepsin specificity, one would expect that incorporation of highly hydrophilic charged residues into a synthetic substrate should not be well accommodated by the enzyme. However, Pohl and Dunn[9] have discovered an extraordinary class of polycationic pepsin substrates, whose kinetic properties

*Scheme 1.* Inhibitor Iva-Val-Val-Sta-Ala-Iaa bound in the active site of porcine pepsin (hydrophobic subsites $S_4$-$S_{3'}$)




contradict the expectation. Although these synthetic peptides contain lysine and arginine residues in positions $P_5$-$P_2$ and $P_{2'}$-$P_{5'}$, they are excellent substrates for pepsin. In some cases the apparent bimolecular rate constant $k_{cat}/K_m$ approaches $10^8$ $M^{-1}$ $sec^{-1}$, a value approximately two orders of magnitude higher than the 'best' hydrophobic substrates. Even more remarkably, the kinetic properties are strikingly dependent on pH. In the case of substrate **17** (Table 1), as the pH is increases by three units, the apparent bimolecular rate constant $k_{cat}/K_m$ increases by more than three orders of magnitude. The Michaelis constant decreases proportionately, while the turnover number is not affected. It was assumed that $K_m$ equals the binding constant for the Michaelis complex, and thus the effects of pH were rationalized in terms of secondary electrostatic interactions with carboxylate groups in the extended active site. This interpretation was supported by indirect crystallographic evidence, based on the studies of *Endothia parasitica* aspartic proteinase[10]. When certain parts of the microbial enzyme sequence were replaced with analogous segments from porcine pepsin, the resulting structural model indicated that several carboxylate residues in the catalytic cleft could engage in productive interactions with a bound cationic ligand, so that the contribution of secondary electrostatic interactions to the increase in binding seemed plausible. Pohl and Dunn[9] also pointed out that nonspecific electrostatic interactions in pepsin catalysis are highly probable. In the primary sequence of 326 residues, the enzyme contains 43 acidic and only 4 basic aminoacids. In fact, the extremely low isoelectric point (pH 1.0) makes porcine pepsin one of the most acidic proteins in nature, and at nonphysiological pH the net charge on an average enzyme molecule is highly negative.

Our aim was to distinguish between the specific and the nonspecific binding models presented for polycationic pepsin substrates by examining the kinetic properties of suitable synthetic inhibitors. We reasoned that for any enzymatic reaction, the inhibition constant ($K_i$) is a less complex function of individual rate constants compared to kinetic parameters of substrates

($k_{cat}$, $k_{cat}/K_m$, $K_m$), and it thus provides a more direct measure of the binding thermodynamics. We report herein the kinetic properties of a series of synthetic inhibitors that contain alkylammonium sidechains in positions $P_4$ through $P_{3'}$, and a statine residue as a replacement for the scissle dipeptidyl subunit ($P_1$-$P_{1'}$). At various values of pH between 2.75 and 5.50, binding to pepsin was compared with the corresponding neutral analogs, which contained alkyl residues of identical length. The results obtained led us to reexamine previously reported kinetic data for polycationic substrates in the context of the Debye-Hückel theory. We also evaluated the effect of pH and ionic strength on the kinetic properties of a polycationic inhibitor. The overall results are consistent with the enhancements of the apparent bimolecular rate constant $k_{cat}/K_m$ arising from nonspecific, long range electrostatic interactions within an early pre-Michaelis complex. These nonspecific interactions are reflected in the kinetics of the corresponding charged inhibitor as a dramatic increase in the inhibitory effect, due to an increase in effective molarity[11] of the inhibitor in the immediate vicinity of each enzyme molecule. Based on our experimental results, we propose a minimal kinetic mechanism for aspartic proteinases which involves 7 distinct enzyme species and 14 primary rate constants.

## Materials and Methods

*Peptide synthesis.* The peptide substrates and inhibitors (Table 1) were synthesized by standard solution phase methods, purified by HPLC and characterized by high resolution FAB-MS and NMR; details will be reported elsewhere[12].

*Enzyme kinetics.* Determination of inhibition constants. For each inhibitor in Table 1, the inhibition constant was first determined at pH 4.0 and ionic strength 100 mM. From a series of experiments with varied inhibitor or varied enzyme concentration, the initial velocities were analyzed by nonlinear fit to Morrison's rate equation for competitive tight binding inhibition[13]. In the determination of the dependence of $K_i$ on pH and ionic strength, a simplified method was used based on equation (1), in which $v_o$ is the rate measured in the absence of the inhibitor. The values of $v_o$ and v were determined in triplicate, at a single concentration of the inhibitor and the enzyme such that velocity v fell within 20 to 80 % of $v_o$. The averaged initial rates were used to compute the inhibition constant directly from equation (1).

$$K_i = \frac{[I]_o - [E]_o (1 - v/v_o)}{(1+[S]_o/K_m) (v_o/v - 1)} \qquad (1)$$

*Quantitative structure - activity correlations.* Hydrophobicity of aminoacids was calculated from atomic increments[14]. Only the sidechain atoms beginning with $C_\beta$ were included in the calculation. Within each series of compounds (**2 - 7** and **8 - 13**), the resulting value of log P for a particular aminoacid was used in the regression analysis for correlating the activity of the whole peptide.

*Table 1* Inhibition of porcine pepsin by pepstatin analogs

| | P5 | P4 | P3 | P2 | P2' | P3' | $K_i$, nM |
|---|---|---|---|---|---|---|---|
| | | Iva | Val | Val | Sta | Ala - Iaa | 0.10 [6] |
| **1** | Boc | -Lys | Val | Val | Sta | Ala - Iaa | 0.22 |
| **2** | | Iva | Lys | Val | Sta | Ala - Iaa | 19.2 |
| **3** | | | Nle | | | | 0.005 |
| **4** | | | Orn | | | | 72.0 |
| **5** | | | Nva | | | | 0.017 |
| **6** | | | Dab | | | | 13.8 |
| **7** | | | Abu | | | | 0.031 |
| **8** | | Iva | Val | Lys | Sta | Ala - Iaa | 6.3 |
| **9** | | | | Nle | | | 0.030 |
| **10** | | | | Orn | | | 7.2 |
| **11** | | | | Nva | | | 0.089 |
| **12** | | | | Dab | | | 21.2 |
| **13** | | | | Abu | | | 0.20 |
| **14** | | Iva | Val | Val | Sta | Lys - Iaa | 0.72 |
| **15** | | Iva | Val | Val | Sta | Ala - Lys-OMe | 0.10 |
| **16** | Lys | Lys | Ala | Lys | Sta | Arg - Leu | 130 |
| **17** | Lys - Lys - Ala - Lys - Phe - Phe(p-$NO_2$) - Arg - Leu | | | | | | --- |

## Results

The peptides shown in Table 1 were prepared and tested in order to examine the effect of positive charge on binding of inhibitors to pepsin. As a parent compound we chose the known[6] pepstatin derived inhibitor, Iva-Val-Val-Sta-Ala-Iaa. The structure was systematically varied by substituting lysine in each position ($P_4$ - $P_{3'}$). In addition, valine residues in $P_2$ and $P_3$ were replaced by other diaminoacids whose alkylamino sidechains

Fig. 1a

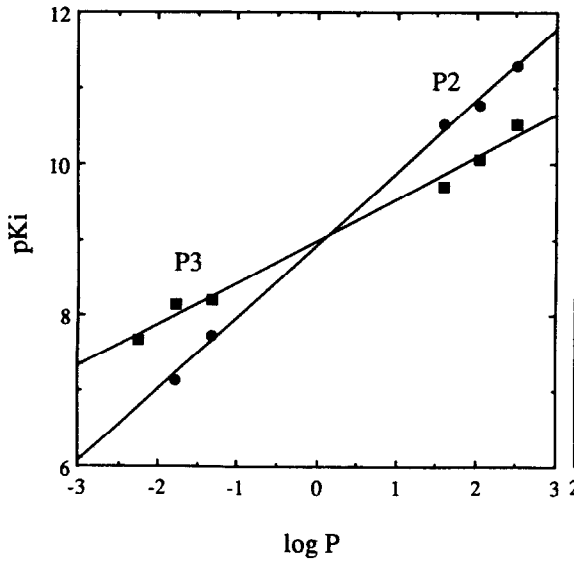
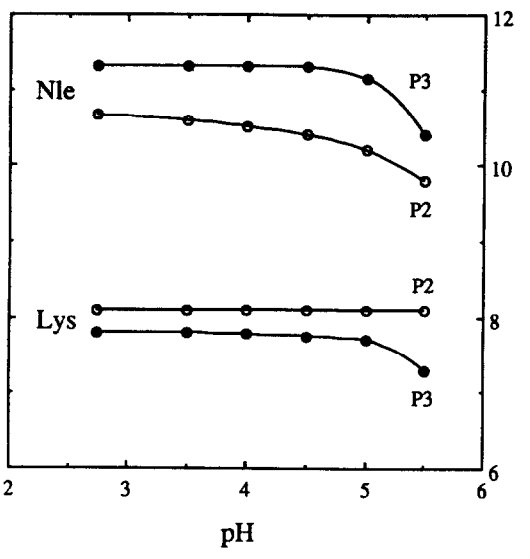



Fig. 1b




contain different numbers of methylene groups. For comparison, desamino analogs with identical alkyl chain length were also prepared. Thus, the lysine (Lys) analogs were designed to be compared with norleucine (Nle), the ornithine (Orn) analogs with norvaline (Nva) and the diaminobutyric acid (Dab) analogs with aminobutyric acid (Abu). The results clearly show that the introduction of positive charge causes an overall decrease in the inhibitory activity. The most significant loss of binding was observed in positions $P_3$ and $P_2$; the inhibition constant changes by a factor of 4000 and 200, respectively, when Lys and Nle analogs are compared. Positions $P_4$ and $P_{2'}$ are less affected, showing a two- and seven-fold decrease in binding upon lysine substitution. Replacement of the C-terminal isoamyl amide residue with lysine methyl ester had no effect. For the two sets of inhibitors with structural variation in $P_2$ or $P_3$, the logarithm of the inhibition constant linearly correlates with the logarithm of the water-octanol partition coefficient (Fig 1a). In terms of changes in free energy, an increase in hydrophobicity produced a proportional increase in binding. The slope in the plot obtained for $P_3$ inhibitors **2** - **7** was +0.95, while $P_2$ substituted analogs **8** - **13** yielded a slope equal to +0.55.

We have determined the effect of pH on the inhibition constants for all lysine substituted inhibitors, as well as for the norleucine control compounds **3** and **9**. The range of experimental pH values was identical with the previous study of polycationic substrates[9]. The results for $P_2$-lysine (**8**) and $P_3$-lysine (**2**) analogs are shown in Fig. 1b. The inhibition constant for **8** showed no change with pH within the whole range of pH 2.75 to 5.50. The

Fig. 2a

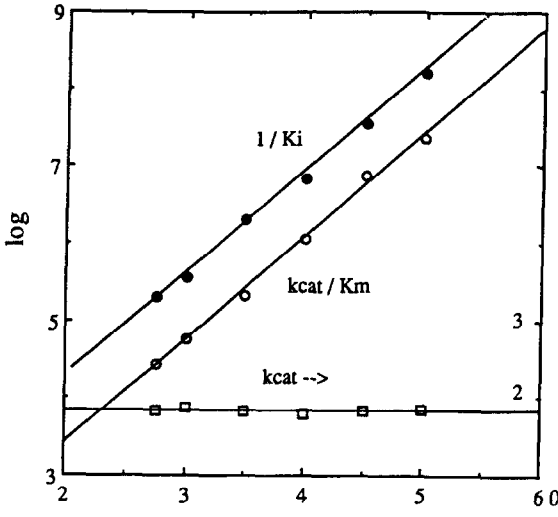
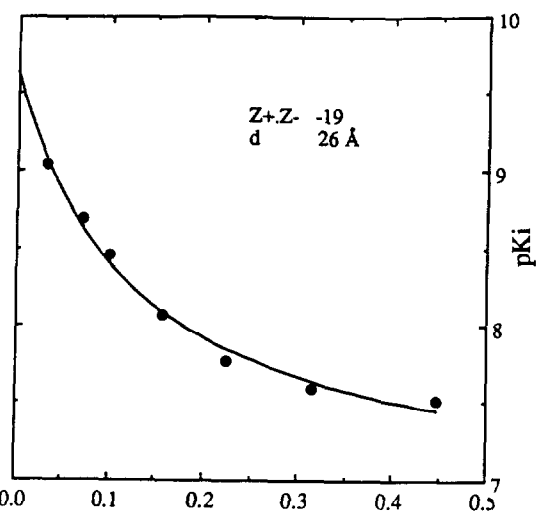



Fig. 2b




binding of **2** decreased slightly above pH 5.00, similar to that found for the norleucine analog **3**. Among all compounds tested, the neutral inhibitor **3** showed the most pronounced loss of binding energy due to the increase in pH; the inhibition constant increased by an order of magnitude from 5 ± 2 pM at pH 4.5, to 45 ± 6 pM at pH 5.5. The other lysine-containing inhibitors have shown a similar lack of sensitivity to pH as shown in Fig. 1b for compounds **2** and **8**. In all $K_i$ determinations mentioned thus far, pepsin was preincubated with the inhibitors before the addition of the substrate, so that the time dependent processes were eliminated. We did focus on possible slow binding in trial experiments. There was a measurably slow onset of inhibition - with a halftime less than a minute - for several charged pepstatin analogs, in accordance with the previous reports on the slow binding of inhibitors based on the propart peptide[15].

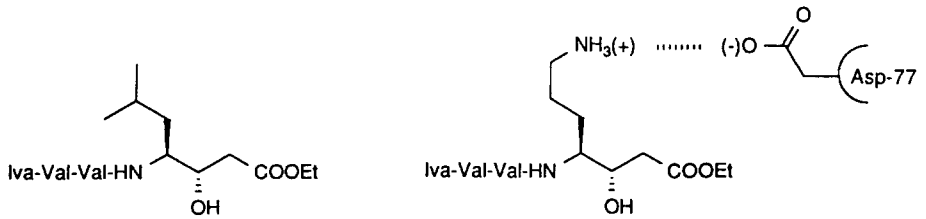
In contrast to the pepstatin analogs that contain a single charged residue, the inhibition constant of the polycationic inhibitor **16** showed a pronounced sensitivity to pH and ionic strength. As the pH increases by 2.2 units (from 2.8 to 5.0), the inhibition increases by three orders of magnitude ($K_i$ decreases from 4.9 μM to 6.3 nM). On the logarithmic scale, the changes in $K_i$ with pH were linear and parallel to the pH dependency of $k_{cat}/K_m$ for the corresponding substrate **17** (Fig. 2a). The substrate kinetic parameters observed by us match the previously reported values[9]. Inhibitor **16** also shows a dramatic dependence of the inhibition constant on the ionic strength of the buffer (Fig. 2b). At pH 4.55, an increase in ionic strength produced a significant loss in inhibitory activity. The inhibition constant

changes from 0.9 nM at low ionic strength (2 mM), to 30 nM at high salt concentration (200 mM). In preliminary experiments without preincubation of the enzyme with the inhibitor, we also monitored the time dependent inhibition. The apparent first order rate constant is strongly dependent on ionic strength; at pH 4.55, the halftime for the formation of the tight complex ranges from 0.6 min (5 mM) to 0.1 min (500 mM).

## Discussion

This study was undertaken to determine whether the pH induced changes in the kinetics of polycationic pepsin substrates[9] are caused by specific electrostatic interactions (ion pair formation in the active site) or by nonspecific, long range ionic association. Our initial approach was to evaluate the effect of alkylammonium substitution on the inhibition constant of pepstatin analogs **1** - **15**. Statine-containing peptides, which are specific inhibitors of aspartic proteinases, have been successfully used as inhibitors of therapeutically important enzymes[16], and to probe the binding subsites in aspartic proteases. We have previously used this approach to identify specific electrostatic interactions in the fungal aspartic proteinase penicillopepsin. Penicillopepsin catalyzes hydrolysis of Lys-Xxx peptide bonds, and from a crystallographic model[17] it was suggested that the lysine can form an ion pair with a carboxylate residue in the active site. The hypothesis was tested with a pepstatin analog[18], in which the isobutyl group in $P_1$ was replaced with butylammonium sidechain to mimic the binding of lysine (Scheme 2). The positively charged LySta-containing inhibitor **19** inhibits penicillopepsin 25 times more strongly than the corresponding neutral analog **18**, which is in accord with the proposed electrostatic binding mode. In a crystallographic study, James et al.[19] later identified the carboxylate in the enzyme subsite $S_1$ as Asp-77.

*Scheme 2*




**18**, $K_i$ 47 nM **19**, $K_i$ 2.1 nM

The specific ion pair model, proposed for binding of polycationic substrates to pepsin,

requires that the positively charged inhibitors show greater affinity for the enzyme, in comparison with the neutral analogs. However, when we compared the inhibitory activity of peptides listed in Table 1, all positively charged pepstatin derivatives were found to be weaker inhibitors than the alkyl substituted compounds. For example, for each pair of inhibitors with identical alkyl chain length in the position $P_3$, the average difference in binding energy in favor of the neutral analog is 2.3 kcal/mol (similarly, 1.5 kcal/mol in $P_2$). Within both groups of inhibitors with systematically varied sidechain length, each additional methylene group causes an increase in binding, which suggests a dominant role of hydrophobic interactions. We have quantitatively established the hydrophobic character in these binding subsites by structure-activity correlations. Using the fragmental sidechain hydrophobicity (log P) as the only molecular descriptor, excellent linear correlations with the logarithmic inhibition constants were obtained for both $P_2$ and $P_3$ analogs (Fig. 1a). Clearly, the interaction energy between enzyme subsites $S_2$, $S_3$ and the inhibitor residues $P_2$, $P_3$ is derived mostly from hydrophobic contacts. The slopes in plots such as in Fig. 1a quantitatively measure the contribution of hydrophobic forces to the total binding energy. In this respect, subsite $P_3$ was found to be more hydrophobic (slope +1.0) than $P_2$ (slope +0.6)

To further characterize the nature of binding interactions in the active site, we determined the pH dependence of inhibition constants for all lysine and norleucine inhibitors. If the specific binding model were operative, an increase in pH should lead to enhanced ion pair formation due to increased ionization of carboxylate residue(s). Consequently, at higher pH the inhibition constant for a lysine-containing inhibitor should decrease. Two conclusions can be drawn from examining the pH profiles in Fig. 1b. First, the binding affinity of the charged inhibitors does not increase with increasing pH, which further testifies against a direct ion pair formation in the active site. Second, above pH 5 there is a decrease in binding for the neutral inhibitors. Since there are no ionizable functional groups in the norleucine-containing inhibitors, these observed changes in binding are likely to result from conformational changes in the enzyme.

The kinetic behavior of compounds **1 - 15** disproves the existence of direct ion pair interactions in the pepsin active site. These results are in accord with most of the previously reported studies of substrate specificity for pepsin, as well as with established structure-activity correlations for pepsin inhibitors, which point towards hydrophobic interactions as the dominant component in ligand binding. How then can we explain the pronounced pH dependence of the bimolecular rate constant $k_{cat}/K_m$ for polycationic pepsin substrates, and the equally dramatic pH sensitivity of the inhibition constant for the corresponding polycationic inhibitors (Fig. 2a)? We propose that the effects are caused by stepwise binding. In the initial diffusion-controlled step, the positively charged ligand is attracted to the surface of the enzyme molecule by long range electrostatic forces. The nonspecific ionic association is then followed by slower surface diffusion into the hydrophobic active site[16]. The

interpretation of experimental kinetic data in terms of stepwise nonspecific binding (Scheme 3) is based on the following arguments.

(1) We invoke the *Debye-Hückel theory* of bimolecular association of ionic reactants, and compare its predictions for steady state kinetic parameters $k_{cat}$ and $k_{cat}/K_m$ with experimental data.

(2) We assume that the 43 carboxylate residues on the enzyme surface have widely overlapping ionization constants. Consequently, *changes in pH produce linear changes in the molecular surface charge ($Z_E$)* within a wide range of experimental pH values.

In Scheme 3, the net rate constant $k_5'$ summarizes all elementary kinetic processes which follow the stepwise binding of the substrate (i.e., chemical and product release steps). The analytical expressions for the turnover number ($k_{cat}$, (2)) and for the apparent bimolecular rate constant ($k_{cat}/K_m$, (3)) were derived by using Cleland's net rate constant method[20]. The polycationic substrate **17** showed no dependence of the turnover number on pH, while the apparent bimolecular rate constant was strongly pH dependent. From this observation and from an inspection of formulas (2) and (3), the pH sensitive step can be tentatively identified. One or more rate constants which appear in equation (3) must be pH dependent, because $k_{cat}/K_m$ is affected by changes in pH. The expression for $k_{cat}/K_m$ contains all rate constant indicated in Scheme 3. However, rate constants which characterize post-binding steps are pH insensitive, because they all appear in $k_{cat}$, and $k_{cat}$ shows no dependence on pH. Thus if we exclude serendipitous mutual compensations, a pH profile such as in Fig. 2a indicates that the strong effect on $k_{cat}/K_m$ is due to pH sensitivity of the initial binding step ($k_1$, $k_2$). This conclusion agrees with the nonspecific binding hypothesis.

*Scheme 3*

$$E + S \underset{k_2}{\overset{k_1}{\rightleftharpoons}} E{\cdot}S \underset{k_4}{\overset{k_3}{\rightleftharpoons}} ES \xrightarrow{k_5'} E + P$$

$$k_{cat} = \frac{k_3\, k_5'}{k_3 + k_4 + k_5'} \qquad (2)$$

$$k_{cat}/K_m = k_1 \frac{k_3\, k_5}{k_2\, k_4 + k_2\, k_5' + k_3\, k_5'} \qquad (3)$$

According to the Debye-Hückel equation (4), the rate of bimolecular association between reactant molecules is determined by molecular charges, and by the ionic strength of the medium. Most relevant to our reinterpretation of the pH dependent substrate kinetics is the linear dependence of the logarithmic bimolecular rate constant $k_1$ on the number of charges, generated on the enzyme surface ($Z_E$). The charge on the ligand is denoted $Z_L$, parameter d is the average interionic distance (Å), I represents the ionic strength (M), and $k_1^\circ$ ($M^{-1}$ $sec^{-1}$) is the limiting value of $k_1$ at zero ionic strength.

$$\log k_1 = \log k_1^\circ + \frac{1.18\, Z_E Z_L \sqrt{I}}{1 + 0.329\, d \sqrt{I}} \qquad (37°C, \text{ water}) \qquad (4)$$

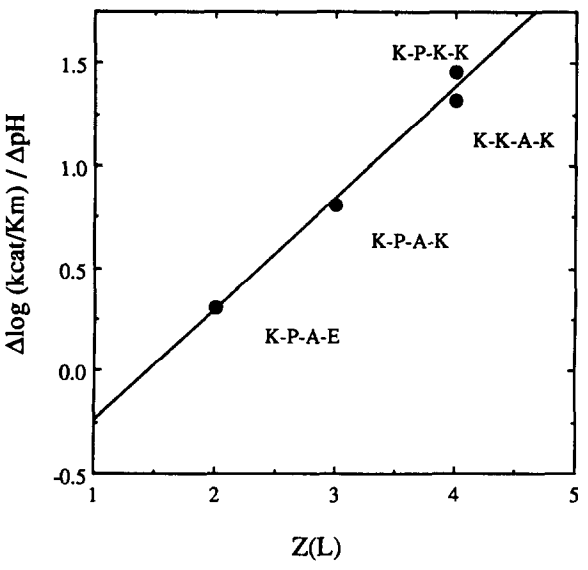
When the pH is increased, the negative surface charge $Z_E$ on an average pepsin molecule will increase linearly, due to ionization of its many carboxylate sidechains. Thus equation (4) predicts a linear increase in log $k_1$ with pH. At the same time, formula (3) establishes proportionality between $k_1$ and $k_{cat}/K_m$, so that the Debye-Hückel theory in fact predicts a *proportional increase in log $k_{cat}/K_m$ with pH*, within a broad range of experimental pH values. Such a pattern was experimentally observed for substrate **17** (Fig. 2a); the prediction based on the nonspecific association model is confirmed. Inherent to our argument is the assumption that the rate constant $k_2$ is not affected by charge effects. In other enzymatic systems with nonspecific electrostatic association[21], it was shown that $k_2$ does to some extent depend on the parameters which appear in equation (4). However the dependence is smaller compared to $k_1$; the overall sensitivity to electrostatic effects was shown to be equivalent for both the true bimolecular rate constant $k_1$, and the apparent bimolecular rate constant $k_{cat}/K_m$ (see Table 2 below).

Compound **17** belongs to a whole series of previously reported pH sensitive substrates[9], in which individual members contain different numbers of lysine and arginine residues; the total positive charge ranges from +2 to +6. We decided to reexamine the kinetic data reported for all these compounds, to establish whether they satisfy the requirements of the nonspecific association model. The analysis was again theoretically based on the Debye-Hückel equation (4), in which $Z_L$ is the positive charge on each particular peptide substrate. In the preceding paragraph we have discussed the linear dependence of log $k_{cat}/K_m$ on pH, which has its theoretical equivalent in the linearity between log $k_1$ and $Z_E$. The slope of such pH dependence - $\partial \log k_1/\partial Z_E$ - is obtained by differentiation of the Debye-Hückel equation with respect to the enzyme surface charge $Z_E$.

$$\frac{\partial \log k_1}{\partial Z_E} = \frac{1.18\, Z_L \sqrt{I}}{1 + 0.329\, d \sqrt{I}} \qquad (5)$$

In the following section we will focus on the effects produced by variations in the molecular

Fig. 3




charge on the ligand. Equation (5) predicts that within a group of substrates, the slopes $\Delta\log(k_{cat}/K_m)/\Delta pH$ obtained from graphs such as in Fig. 2a, should be linearly dependent on the number of basic aminoacids in each substrate's sequence ($Z_L$). Moreover, it is known that nonspecific ionic association originates in long range electrostatic forces (10 - 20 Å), so that the slopes in graphs of $\log(k_{cat}/K_m)$ *vs.* pH should *not* depend on the exact position of basic aminoacids in the sequence (the lysine residues in each substrate are not expected to engage in ionic interactions with any particular carboxylate on the enzyme surface). The results of our analysis for a group of four pepsin substrates with a general structure $A_1$-$A_2$-$A_3$-$A_4$-Phe-Phe(p-$NO_2$)-Arg-Leu are shown in Fig. 3 (one-letter aminoacid code is used to identify the N-terminal fragment of the substrates). For each substrate, we have first determined the slope[22] in the graph of pH versus $\log(k_{cat}/K_m)$ from published data[9], and then plotted the results against the total number of positive charges in the substrates ($Z_L$). The predictions of the Debye-Hückel theory are satisfied in both respects. The slopes $\Delta\log(k_{cat}/K_m)/\Delta pH$ do linearly depend on the number of basic aminoacids. They do *not* depend on their exact position, as is exemplified in substrates with N-terminal sequences Lys-Lys-Ala-Lys and Lys-Pro-Lys-Lys.

So far we have supported the evidence for long range, two step ionic binding of polycationic pepsin substrates only by semiquantitative arguments. However, the Debye-Hückel theory on which they were based also represents a convenient quantitative model. Nolte and collaborators[21] used it in a series of experiments with varied ionic strength, to determine the number of charges on the surface of acetylcholine esterase which are engaged in binding of the monocationic substrate. It was shown that both the apparent bimolecular rate constant $k_{cat}/K_m$, and the true association rate constant $k_1$, depended on the ionic strength to approximately the same extent. When the experimental data for either $k_{cat}/K_m$ or $k_1$ were fitted to a linearized form of eq. (5), similar estimates for the enzyme surface charge ($Z_E$) and the average interionic distance (d) were obtained. These authors also determined the effect of ionic strength on the ligand binding constant $K_d$. As the ionic strength of the buffer increased, the binding of the charged ligand decreased. We have analyzed the published

binding data by nonlinear least squares fit (Marquardt algorithm), and compared the estimated values of $Z_E$ and d with the results for the two methods indicated above (based on $k_1$ and $k_{cat}/K_m$, see Table 2).

Because the dissociation rate constant $k_2$ to some extent does depend on the ionic strength, the estimates of $Z_E$ and d obtained from the rate data ($k_1$) and from the equilibrium data ($K_d$) are somewhat different. The analysis based on $K_d$ tends to overestimate both in the number of charges, and in the average interionic distance. Nevertheless, all three methods indicated in Table 2 ($k_1$, $k_{cat}/K_m$, or $K_d$) convey the same message: from the average distance of about 10 Å, the monocationic ligands can kinetically 'see' approximately 7 negative charges on the enzyme surface.

*Table 2* Ionic strength dependent kinetics of acetylcholine esterase[21]

| observable | $Z_E$ | d (Å) |
|---|---|---|
| $k_1$ | 6.3 | 9 |
| $k_{cat}/K_m$ | 8.5 ± 2.2 | 12 ± 3 |
| $K_d$ | 10 ± 3 | 14 ± 5 |

Encouraged by the successful analysis of the equilibrium binding data reported by Nolte et al.[21], we applied the same approach to a polycationic pepsin inhibitor, Lys-Lys-Ala-Lys-Sta-Arg-Leu (**16**). Its structure was derived from the pH sensitive substrate, Lys-Lys-Ala-Lys-Phe-Phe($NO_2$)-Arg-Leu (**17**), by replacing the scissle unit -Phe-Phe($NO_2$)- with statine (Sta), in accord with the established precedent that this non-proteinogenic hydroxyaminoacid acts as a dipeptidyl transition state analog for aspartic proteinases[16]. We intended to use the inhibition constant $K_i$ as a convenient surrogate for the bimolecular rate constant $k_{cat}/K_m$. For this purpose, it was necessary to compare the sensitivity of $K_i$ and $k_{cat}/K_m$ to charge effects, which was most conveniently accomplished by examining the variations in both steady state parameters due to changes in pH (Fig.2a). There is a close parallelism in pH-induced changes in the inhibition constant of Lys-Lys-Ala-Lys-Sta-Arg-Leu, and the apparent bimolecular rate constant of the corresponding substrate. A replot of $pK_i$ vs. $\log(k_{cat}/K_m)$ yields a straight line with a slope of 0.96 and correlation coefficient 0.97. These results establish that $K_i$ and $k_{cat}/K_m$ react identically to physical factors appearing in the Debye-Hückel equation (in the case of variations in pH, the corresponding parameter is the enzyme surface charge $Z_E$).

Finally, we determined the effect of ionic strength on the inhibition constant. The initial velocities were analyzed by nonlinear least squares fit to the Debye-Hückel equation (5), and the results are shown in Fig 2b. The estimated average interionic distance is 26 ± 8 Å, in agreement with the physical model for nonspecific (long range) electrostatic association. The Coulombic product $Z_E Z_L$ is -19 ± 6. If we assume that all four positive charges in the

substrate participate in the binding, then the average number of carboxylate residues on the pepsin surface with which they interact is five ($Z_E Z_L$ = 4 x 5). Other possible combinations are 3 x 6 and 2 x 9. In conjunction with the kinetic data obtained for pepstatin analogs **1-15**, which were used here as electrostatic active site probes, these results show that the steady state kinetic parameters for polycationic pepsin substrates as well as for inhibitors are strongly influenced by long range electrostatic interactions.

## Minimal Kinetic Mechanism for Catalysis by Aspartic Proteinases

*Scheme 4*

$$E \underset{k_2}{\overset{k_1}{\rightleftharpoons}} E{\cdot}S \underset{k_4}{\overset{k_3}{\rightleftharpoons}} ES \underset{k_6}{\overset{k_5}{\rightleftharpoons}} EPQ \underset{k_8}{\overset{k_7}{\rightleftharpoons}} EP{\cdot}Q \xrightarrow{k_9} EP \underset{k_{12}}{\overset{k_{11}}{\rightleftharpoons}} E{\cdot}P \xrightarrow{k_{13}} E$$

$$k_{cat} = \frac{k_3\, k_5\, k_7\, k_9\, k_{11}\, k_{13}}{\begin{aligned} &k_4\, k_6\, k_8\, k_{11}\, k_{13} + k_4\, k_6\, k_9\, k_{11}\, k_{13} + k_4\, k_7\, k_9\, k_{11}\, k_{13} + k_5\, k_7\, k_9\, k_{11}\, k_{13} + \\ &k_3\, k_6\, k_8\, k_{11}\, k_{13} + k_3\, k_6\, k_9\, k_{11}\, k_{13} + k_3\, k_7\, k_9\, k_{11}\, k_{13} + k_3\, k_5\, k_8\, k_{11}\, k_{13} + \\ &k_3\, k_5\, k_9\, k_{11}\, k_{13} + k_3\, k_5\, k_7\, k_{11}\, k_{13} + k_3\, k_5\, k_7\, k_9\, k_{11} + k_3\, k_5\, k_7\, k_9\, k_{12} + \\ &k_3\, k_5\, k_7\, k_9\, k_{13} \end{aligned}}$$

$$\frac{k_{cat}}{K_m} = \frac{k_1\, k_3\, k_5\, k_7\, k_9}{k_2\, k_4\, k_6\, k_8 + k_2\, k_4\, k_6\, k_9 + k_2\, k_4\, k_7\, k_9 + k_2\, k_5\, k_7\, k_9 + k_3\, k_5\, k_7\, k_9}$$

We have presented two lines of evidence for long range association of pepsin with its polycationic substrates as a kinetically significant process which precedes the binding in the active site. Firstly, we reanalyzed the published substrate data. The long range association model is supported by the linear dependence of log $k_{cat}/K_m$ on pH within several pH units, by the linear dependence of slopes in the corresponding graphs on the number of lysines residues in substrate sequence, and by the insensitivity of such slopes to the exact position of charged aminoacids. Secondly, we studied the binding of a polycationic inhibitor **17**. We monitored the two-step binding directly, by following the slow onset of inhibition, and established the physical parameters for long range association (average interionic distance 26 Å). The involvement of stepwise binding has important implications for the minimal kinetic

mechanism, which has to be considered for pepsin and other aspartic proteinases. If the substrate and the enzyme initially form a loosely associated ionic complex E·S, then we need to postulate the existence of analogous complexes also on the part of the products (EP·Q, E·Q), because the essential ionic character of ligands is preserved in the hydrolytic reaction. Moreover, molecular aggregation due to forces other than electrostatic can be involved in nonspecific binding; it has been demonstrated that certain hydrophobic substrates of pepsin have 4 to 5 binding sites on the enzyme surface[23]. The minimal kinetic mechanism in Scheme 4 thus reflects the effects of nonspecific binding in general.

The mechanism proposed here contains seven elementary kinetic steps. Rich and Northrop[24] proposed a similar six-step mechanism from a different perspective, by considering the conformational changes in the enzyme upon binding of substrates. Crystallographic evidence[25] shows that the binding of competitive inhibitors to aspartic proteinases is accompanied by pronounced movement of a domain in the enzyme tertiary structure. This 'flap' region closes upon the inhibitor and effectively entraps it in the active site. It is very likely that binding of substrates occurs in the same fashion, and that the full catalytic sequence also contains conformationally trapped forms of ES, EPQ, and EP. If these were inserted into Scheme 4, we would obtain a kinetic mechanism which involves 10 enzyme species and 20 primary rate constants. The significance of properly considering such complex kinetic sequences, particularly with regard to the kinetics of protease inhibitors, is discussed in detail in the work cited above[24].

## Conclusion

The nonspecific electrostatic interactions in pepsin catalysis are very significant. An adequate minimal kinetic mechanism involves seven distinct enzyme species. With this mechanism, we can successfully interpret the structure-activity correlations for charged pepsin substrates and inhibitors, as well the dependence of their steady state parameters on physical factors. An observation which deserves special attention is the dramatic pH and ionic strength dependence of $K_i$ for the polycationic inhibitor. If we interpret the pH and ionic strength effects on the kinetics of charged substrates by *long range* electrostatic interactions - as opposed to active site ion pair formation - then why do these physical factors so strongly influence the inhibition constant? Intuitively it could be expected that $K_i$ reflects only the thermodynamics of intimate enzyme-inhibitor interactions, and not the nonspecific long range effects. The assumption that $K_i$ represents a thermodynamic equilibrium constant is commonly applied to inhibitors acting as ground state as well as transition state analogs.

With the involvement of nonspecific enzyme-ligand interactions, we encounter an essential difference between substrates and inhibitors, which provides an explanation for the apparent

discrepancy. In both cases, the nonspecific interactions (ionic, hydrophobic, or van der Waals) tend to increase the effective molarity[11] of the ligand in close vicinity of an enzyme molecule. In the case of a substrate, the microscopic concentration does not necessarily increase because the substrate is being consumed, so that the net effect is an increase in the bimolecular association rate constant. In the theory of the diffusion control in enzymatic reactions, this effect of nonspecific interactions is predicted by Chou's model[26]. On the other hand, in the case of an inhibitor, the local concentration near the enzyme molecule does increase due to nonspecific binding. Consequently, the inhibitory effect is higher, when compared to a hypothetical case of uniformly distributed inhibitor concentration[27].

## Acknowledgements

Thanks are due to Dr. Kuo-Chen Chou, The Upjohn Co., for valuable comments on the manuscript. Financial support by the National Institutes of Health (DK20100) is gratefully acknowledged.

## References and Notes

1. *Primary specifity* of proteases denotes a preference of the enzyme for a particular aminoacid as a part of the peptide bond being hydrolyzed. The terms *secondary specifity* and *secondary interactions* refer to kinetic effects produced by aminoacids at some distance from the cleaved peptide bond.
2. Fruton, J. *Adv. Enzymol. Related Areas Mol. Biol.***1976**, *44*, 1-36.
3. Schlamowitz, M.; Shaw, A.; Jackson, W.T. *J. Biol. Chem.* **1968**, *243*, 2821-2828.
4. Tang, J. *J. Biol. Chem.* **1965**, *240*, 3810-3815.
5. Shin, H.D. *Bull. Chem. Soc. Japan* **1970**, *43*, 3472-3474.
6. Rich, D.H.; Bernatowicz, M.S. *J. Med. Chem.* **1982**, *25*, 791-795.
7. Schechter, I.; Berger, A. *Biochem. Biophys. Res. Commun.* **1967**, *27*, 157-162 .
8. *Specific binding, specific interactions* refer to the binding of ligands in the enzyme active site, while the terms *nonspecific binding, nonspecific interactions* indicate that binding occurs outside the active site.
9. Pohl, J.; Dunn, B.M. *Biochemistry* **1988**, *27*, 4827-4834.
10. Foundling, S.I.; Cooper, J.; Watson, F.E.; Cleasby, A.; Pearl, L.H.; Sibanda, B.L.; Hemmings, A.; Wood, S.P.; Blundell, T.L.; Valler, M.J.; Norey, C.G.; Kay, J.; Boger, J.; Dunn, B.M.; Leckie, B.J.; Jones, D.M.; Atrash, B.; Hallet, A.; Szelke, M. *Nature (London)* **1987**, *327*, 349-352.
11. [a] Haberfield, P.; Cincotta, J.J. *J.Org.Chem.* **55**, 1334-1338 (1990). [b] Kirby, A.J. Adv.Phys.Org.Chem. **17**, 183 (1980).

12. Kuzmic, P.; Sun, C.-Q.; Zhao, Z.-C.; Rich, D.H. manuscript in preparation.
13. Morrison, J.F. *Biochim. Biophys. Acta* **1969**, *185*, 269-286.
14. [a] Ghose, A.K.; Crippen, G.M. *J. Comput. Chem.* **1986**, *7*, 565-577. [b] Furet, P.; Sele, A.; Cohen, N.C. *J. Mol. Graphics* **1988**, *6*, 182-189. [c] Eisenberg, D.; McLachlan, A.D. *Nature (London)* **1986**, *319*, 199-203.
15. Dunn, B.M.; Parten, B.; Jimenez, M.; Rolph, C.E.; Valler, M.; Kay, J. in *"Aspartic Proteinases and their Inhibitors"* (V. Kostka, Ed.) Walter de Gruyter, Berlin, pp. 221-243 (1985).
16. [a] Rich, D.H. in *"Proteinase Inhibitors"* (A.J. Barrett and G. Salveson, Eds.), Elsevier, New York, pp. 179-217 (1986). [b] Rich, D.H. in *"Comprehensive Medicinal Chemistry, Vol. 2"* (P.G. Sammes, Ed.), Pergamon Press, Oxford, pp. 391-443 (1990).
17. James, M.N.G.; Sielecki, A.; Salituro, F.; Rich, D.H.; Hofmann, T. *Proc. Natl. Acad. Sci. USA* **1982**, *79*, 6137-6141.
18. [a] Rich, D.H.; Salituro, F.G.; Holladay, M.W. in *"Peptides, Structure and Function"*, Proceedings of the 8th American Peptide Symposium (V. Hruby, D. H. Rich, Eds.), Pierce Chemical Co.; Rockford, Illinois, pp. 511-520 (1983). [b] Salituro, F.C.; Agarwal, N.; Hofmann, T.; Rich, D.H. *J. Med. Chem.* **1987**, *30*, 286-295.
19. James, M.N.G.; Sielecki, A.R.; Hofmann, T. in *"Aspartic Proteinases and their Inhibitors"* (V. Kostka, Ed.), Walter de Gruyter, Berlin, pp. 163-177 (1985).
20. Cleland, W.W. *Biochemistry* **1975**, *14*, 3220-3224.
21. Nolte, H.J.; Rosenberry, T.L.; Neumann, E. *Biochemistry* **1980**, *19*, 3705-3711.
22. A complicating factor in our analysis was that most of the polycationic substrates reported in the literature showed rather complex pH profiles, compared to substrate **17**. A typical feature was a break in the slope of $\log(k_{cat}/K_m)$ in the high pH region. In light of the abrupt changes in the inhibition constant detected for the neutral inhibitor **3** above pH 5.0, which we interpret as a conformational change in the enzyme, we have disregarded data falling into the corresponding region.
23. Raju, E.V.; Humphreys, R.E.; Fruton, J.S. Biochemistry **1972**, *11*,3533-3535.
24. Rich, D.H.; Northrop, D.B. in *"Computer Aided Drug Design"* (T.J. Perun and C.L. Prost, Eds.), Marcel Dekker, New York, p. 185-250 (1989).
25. Miller, M.; Schneider, J.; Sathyanarayna, B.K.; Toth, M.V.; Marshall, G.R.; Clawson, L.; Selk L.; Kent, S.B.H.; Wlodawer, A. *Science (Washington, D.C.)* **1989**, *246*, 1149-1152.
26. [a] Chou, K.C.; Jiang, S.P. *Sci. Sin.* **1974**, *17*, 664-680. [b] Chou, K.C. *Sci. Sin.* **1976**, *19*, 505-528. [c] Chou K.C., Forsén, S. *Biophys. Chem.* **1980**, *12*, 255-263. [d] Chou, K.C.; Zhou, G.P. *J. Amer. Chem. Soc.* **1982**, *104*, 1409-1413.
27. A referee has suggested that the observed effects of pH on $K_i$ and $k_{cat}/K_m$ could be explained by the deprotonation of a lysine sidechain with drastically lowered $pK_a$ (normally, $pK_a$ 10.53). However, pH titration experiments excluded that possibility.