Pear, M. R., & Weiner, J. H. (1979) *J. Chem. Phys. 71*, 212.
Quentrec, N., & Brot, C. (1973) *J. Comput. Phys. 13*, 430.
Rabolt, J. F., Moore, W. H., & Krimm, S. (1977) *Macromolecules 10*, 1065.
Richards, F. M. (1977) *Annu. Rev. Biophys. Bioeng. 6*, 151.
Rigler, R. (1977) *Trends Biochem. Sci. 2*, 252.
Rossky, P. J., & Karplus, M. (1979) *J. Am. Chem. Soc. 101*, 1913.
Ryckaert, J.-P., Ciccotti, G., & Berendsen, H. J. C. (1977) *J. Comput. Phys. 23*, 327.
Saviotti, M. L., & Galley, W. C. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71*, 4154.
Shepherd, I. W. (1975) *Rep. Prog. Phys. 38*, 565.
Shimanouchi, T. (1976) *Structural Studies of Macromolecules by Spectroscopic Methods* (Ivin, K. J., Ed.) Wiley, New York.
Stillinger, F. H., & Rahman, A. (1972) *J. Chem. Phys. 57*, 1281.
Suezaki, Y., & Go, N. (1975) *Int. J. Pept. Protein Res. 7*, 333.
van Gunsteren, W. F. (1980) *Mol. Phys. 40*, 1015.
van Gunsteren, W. F., & Berendsen, H. J. C. (1977) *Mol. Phys. 34*, 1311.
van Gunsteren, W. F., & Karplus, M. (1980) *J. Comput. Chem. 1*, 266.
Verlet, L. (1967) *Phys. Rev. 159*, 98.
Wagner, G., & Wüthrich, K. (1978) *Nature (London) 275*, 247.
Weber, T. A., & Helfand, E. (1979) *J. Chem. Phys. 71*, 4760.
Woodward, C. K., & Hilton, B. D. (1979) *Annu. Rev. Biophys. Bioeng. 8*, 99.
Yu, N.-T. (1977) *CRC Crit. Rev. Biochem. 4*, 229.

Thermodynamics and Kinetics of Single Residue Replacements in Avian Ovomucoid Third Domains: Effect on Inhibitor Interactions with Serine Proteinases†

Mark W. Empie and Michael Laskowski, Jr.*

ABSTRACT: Sequence determinations in our laboratory have yielded the primary structures of ovomucoid third domains from 35 avian species. From this list, 12 sequences could be arranged into a contiguous set such that each sequence differs from a second by a single amino acid replacement. For this set of domains and for five additional domains of special interest, we report here the association equilibrium constants for their binding with bovine α-chymotrypsin, elastase I, and subtilisin Carlsberg. The results are interpreted with the aid of the three-dimensional structure of highly homologous Japanese quail ovomucoid third domain and of computer-generated models of the complexes of the inhibitor with the respective enzymes. The results show that (i) changes in inhibitor residues other than the primary recognition residue (P_1), even sequentially far from the reactive site, may exert large effects on association equilibrium constant values provided these residues make contact with the enzyme, (ii) changes in residues other than P_1 often exert large differential effects toward the different enzymes, i.e., the same change can make the inhibitor stronger for one enzyme and weaker for another, (iii) the sign and to some extent the magnitude of the changes can be rationalized from the known structures of the inhibitor and the enzyme, (iv) changes in surface residues which do not contact the enzyme in complex are virtually without effect, and (v) glycosylated and nonglycosylated inhibitors have the same constants. For confirmation of the validity of the equilibrium constant comparisons in a few cases, the rate constants k_{on} and k_d were determined and the resultant calculated equilibrium constant values compared to the directly determined numbers. An additional test of validity is provided by experiments where a glycosylated domain of one species is allowed to compete with an unglycosylated domain of another for the same enzyme.

It is a major objective of this laboratory to construct an algorithm which would predict, from only the amino acid sequence, the thermodynamic and kinetic parameters for the interaction of serine proteinases with their protein inhibitors (Laskowski, 1980; Laskowski et al., 1981). The kinetic and thermodynamic properties are known for a number of these protein inhibitors which follow an analogous, standard mechanism (Laskowski & Sealock, 1971; Laskowski & Kato, 1980). However, not all these inhibitors are homologous; instead, they can be grouped into several families (Laskowski & Kato, 1980). In development of the algorithm, it is essential to focus first on a single family. To date, systematic studies within a family have been limited and have focused on the two residues, P_1 and P_1' (Schecter & Berger, 1967), which make up the reactive site peptide bond (Sealock & Laskowski, 1969; Kowalski et al., 1974; Kowalski & Laskowski, 1976; Odani & Ikenaka, 1978; Odani & Ono, 1980; Wenzel & Tschesche, 1981). For a variety of reasons, but mainly the availability of needed, naturally occurring variants, we have chosen to focus on the Kazal family of inhibitors. In particular, this paper concentrates on a subset of this family—the third domain of avian ovomucoid.

Ovomucoid is a major component (~10% of the protein) of avian egg whites and is responsible for most of the inhibitory activity against serine proteinases in the egg white. Intact ovomucoid was found by Feeney and co-workers (Rhodes et al., 1960; Feeney, 1971) to have two startling characteristics: (1) a single ovomucoid molecule, depending on the avian

† From the Department of Chemistry, Purdue University, West Lafayette, Indiana 47907. *Received November 13, 1981.* Supported by Grant GM 10831 from the National Institutes of General Medical Sciences, National Institutes of Health. Many of the results given here were presented in preliminary form (Empie & Laskowski, 1981; Laskowski et al., 1981).

0006-2960/82/0421-2274$01.25/0

species, may simultaneously combine with one, two, or three enzyme molecules (i.e., they may be single headed, double headed, or triple headed), and (2) ovomucoids from closely related species have strikingly different inhibitory properties. Both of these findings were rationalized when it was shown that ovomucoids consist of three homologous tandem domains; each domain is potentially capable of binding a serine proteinase (Kato et al., 1976) at its unique reactive site peptide bond. The third domain can be cleaved from the other two by limited proteolysis (Kato et al., 1978). This particular domain is heterogeneously glycosylated, and one may isolate by gel permeation chromatography both the glycosylated and nonglycosylated forms. A large number of third domains from various avian species have been isolated and sequenced (Kato et al., 1978). From this list, 12 sequences have been selected such that pairs can be formed which contain only a single amino acid substitution between the two members. An additional set of five can be found which contain only a few amino acid changes to the nearest sequence. The effect of these amino acid substitutions in native ovomucoid third domain on the association equilibrium constant, K_a, in a few cases the association rate constant, k_{on}, and the dissociation rate constant, k_d, is compared for various enzymes.

The mechanism of interaction for these and a great number of other inhibitors with serine proteinases was first put forth by Luthy et al. (1973). It can be given by[1]

$$E + I \underset{}{\overset{K_L}{\rightleftharpoons}} L \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} C \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} L^* \underset{K_{L^*}}{\rightleftharpoons} E + I^*$$

where E is the proteinase, I and I* are the virgin and modified (reactive site peptide bond hydrolyzed) inhibitor, respectively, C is the stable complex, and L and L* are loose, noncovalent complexes of enzyme with virgin or modified inhibitor. Complexes L and L* are in extremely rapid equilibrium with the free enzyme and inhibitors. At low enzyme and inhibitor concentrations, species L and L* are insignificant and the mechanism reduces to

$$E + I \underset{k_{-2}}{\overset{k_2/K_L}{\rightleftharpoons}} C \underset{k_{-3}/K_{L^*}}{\overset{k_3}{\rightleftharpoons}} E + I^*$$

where k_2/K_L and k_{-3}/K_{L^*} are the second-order rate constants k_{on} and k_{on}* for the association of enzyme with virgin and modified inhibitor, respectively. The constants k_{-2} and k_3 become k_{off} and k_{off}*, respectively. For a number of enzymes, the rate constants for the modified inhibitor interaction are slow. Hence, practical thermodynamic and kinetic data using virgin inhibitor may involve only the reactions up to complex and no further conversion to modified inhibitor. Hence, k_{off} is much greater than k_{off}* so that k_d (defined as $k_d = k_{off} + k_{off}$*) is approximately equal to k_{off}. This is true for trypsin and chymotrypsin (and probably elastase) but is not general for all serine proteinases. W. Ardelt and M. Laskowski, Jr. (unpublished results), have shown that k_{off} and k_{off}* have similar values in the dissociation of turkey ovomucoid third domain–subtilisin complex. Thus, in the k_d value for this system, both k_{off} and k_{off}* are significant.

Collection of binding data in the older literature has been hampered by the lack of a routine, facile method of measurement for tight binding inhibitors. Thus, much of the literature data is for weak inhibitors or for binding at pH values far from optimum (Laskowski & Sealock, 1971). Chemical modification studies to probe the effect of specific side chains have also, generally, only detected grossly weakened inhibition (Laskowski & Sealock, 1971). Quantitation of subtle differences in binding is rare.

This problem results from the absence of an accurate, general method for measuring tight binding equilibrium constants (K_a's > 10^8 M^{-1}). The method of Green & Work (1953) is applicable for these determinations but suffers from a number of difficulties (Laskowski & Sealock, 1971; Bieth, 1974). The principal problem is that enzyme and inhibitor concentrations must be sufficiently low so that measurable amounts of enzyme remain uncomplexed after the enzyme–inhibitor equilibrium has been attained. The recent development of *p*-nitroanilide, extended peptide substrates which have high k_{cat}/K_m values (Del Mar et al., 1979; Bieth et al., 1974; Pozsgay et al., 1979) and stable, highly sensitive spectrophotometers allow one to accurately measure extremely low levels of free enzyme. This permits facile measurement of binding constants as high as 10^{12} M^{-1} by the Green & Work (1953) method; computer fitting by nonlinear least-squares analysis of all the data points in the titration curve greatly increases precision.

The results from this paper are intended to shed light on the nature of serine proteinase–protein inhibitor interactions and what is required to be a good Kazal inhibitor. These requirements have been addressed for the inhibitor mechanism in general, but the role of specific sequence changes in homologous inhibitors toward inhibitory parameters has not. This report is the beginning of such a study.

Experimental Procedures

Materials

Nonfertile avian eggs from various avian species were obtained fresh from domestic game breeders. Bovine trypsin and α-chymotrypsin A were purchased from Worthington Biochemical Corp. Elastase I, 3 times crystallized, was a gift of M. Laskowski, Sr., and subtilisin Carlsberg was purchased from Sigma Chemical Co.

The chromogenic turnover substrates *N*-succinyl-L-alanyl-L-alanyl-L-prolyl-L-phenylalanine *p*-nitroanilide, *N*-succinylglycylglycyl-L-phenylalanine *p*-nitroanilide, *N*-succinyl-L-alanyl-L-alanyl-L-alanine *p*-nitroanilide, and *N*-benzoyl-L-valylglycyl-L-arginine *p*-nitroanilide were products of Vega Biochemical. The trypsin-specific burst titrant *p*-nitrophenyl *p'*-guanidinobenzoate was a product of Nutritional Biochemicals. The fluorescent burst titrant 4-methylumbelliferyl *p*-guanidinobenzoate (Jameson et al., 1973) was synthesized in this laboratory by Dr. P. Fankhauser, using methods similar to those of Chase & Shaw (1967) and Wang & Shaw (1972) for the synthesis of the analogous *p*-nitrophenyl esters, and the fluorescent burst titrant 4-methylumbelliferyl *p*-(*N*,*N*,*N*-trimethylammonio)cinnamate was purchased from Sigma Chemical Co. All substrates were dissolved in dimethyl sulfoxide previously dried over molecular sieves.

Tris-HCl[2] buffer and Triton X-100 were obtained from Sigma Chemical Co. All other chemicals used were reagent grade or the best commercially available.

Methods

Preparation of Ovomucoid Third Domain. Entire ovomucoid was isolated by the modified procedure of Lineweaver & Murray (1947) as given in Bogard et al. (1980). Ovomucoid third domain was prepared by limited proteolysis using *Staphylococcus aureus* V-8 proteinase (Kato et al., 1978) and

[1] We agree that the even more extended mechanism including an additional intermediate, X, on the modified side and first proposed by Quast et al. (1978) is probably more correct, but such an elaborate mechanism is irrelevant to the considerations made here.

[2] Abbreviations: Tris, tris(hydroxymethyl)aminomethane; K_a, association equilibrium constant.

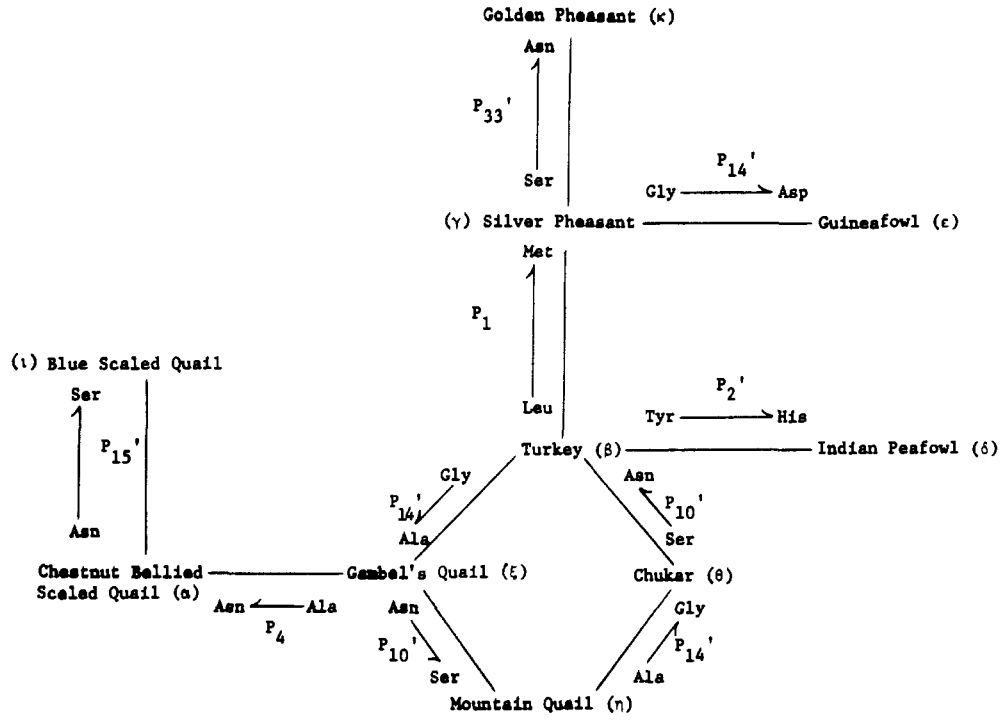

FIGURE 1: Ten sequences listed in Table IA can be contiguously related by single amino acid substitutions. The related sequences are listed by species name and sequence designator. The position and the substituted amino acids are also given. Note that in a few cases the single amino acid difference is true only if an additional difference in the first five NH_2-terminal residues is neglected (see Table IA and Discussion).

purified by the method of Kato.

Purification of Ovomucoid Third Domains. A final purification step was made by diethylaminoethyl-Sepharose ion-exchange chromatography at pH 8 with a 0–0.5 M sodium chloride gradient elution. This final step is required to ensure homogeneity of the sample by removing any modified inhibitor, inhibitor which had been nicked during the limited proteolysis or contaminating material arising from deamidation. Possible deamidation is obviously serious for this study since it not only creates a single residue replacement but also changes a neutral residue into a nonconservative charged residue; nicking is well-known to dramatically change the inhibitor conformation and to weaken inhibitory activity.

Amino Acid Analysis. Amino acid analysis was carried out on all isolated ovomucoid third domains. Protein samples were hydrolyzed in vacuo in sealed 9-mm tubes which had been previously cleaned in 6 N HCl. Protein concentrations were 0.2 mg in 0.75 mL of constant boiling HCl. Following hydrolysis for 18–24 h, the sealed ampules were opened and dried under vacuum. Analysis was carried out after redissolution on a Durrum D-500 amino acid analyzer. A computer integer fit program after Hoy et al. (1974) calculated the results.

Peptide Analysis. Two ovomucoid third domain samples, from blue scaled quail and chestnut bellied scaled quail, were suspected of being polymorphic at the residue position of their single amino acid difference. The peptide containing this position was isolated from individual eggs and analyzed by the method of Bogard et al. (1980) to ensure the absence of polymorphic materials in the test samples.

Stock Enzyme and Inhibitor Solutions. α-Chymotrypsin and elastase I stock solutions were made by dissolving protein in millimolar HCl containing 0.02 M $CaCl_2$. Subtilisin stock was made in pH 4.5 acetate buffer. Final stock concentrations for both enzymes and inhibitors (also in millimolar HCl) were 10^{-4} M.

Enzyme Assays. All initial rate assays for α-chymotrypsin, elastase I, and subtilisin Carlsberg were carried out by using the commercially available turnover substrates. The production of *p*-nitroaniline was measured at a wavelength of 410 nm as a function of time by a Cary 118 spectrophotometer, using 1-cm path-length quartz cuvettes and 3-mL assay volume. The buffer in all cases was 0.2 M Tris and 0.02 M $CaCl_2$ (pH 8.30) with 0.005% (w/v) Triton X-100 added. α-Chymotrypsin assays were carried out by using 10^{-7} M enzyme and *N*-succinylglycylglycyl-L-phenylalanine *p*-nitroanilide (7.8×10^{-5} M) as substrate (Shinar & Gertler, 1979) or enzyme at 10^{-10} M concentration and *N*-succinyl-L-alanyl-L-alanyl-L-prolyl-L-phenylalanine *p*-nitroanilide as substrate (5.3×10^{-5} M) (Del Mar et al., 1979). Elastase assays were performed at 10^{-8} M and lower concentrations using *N*-succinyl-L-alanyl-L-alanyl-L-alanine *p*-nitroanilide (5.3×10^{-4} M; Bieth et al., 1974). Subtilisin assays employed the Shinar substrate given above for enzyme levels at 10^{-7} M and *N*-benzoyl-L-valylglycyl-L-arginine *p*-nitroanilide (2.7×10^{-4} M) for enzyme levels of 10^{-10} M.

Enzyme and Inhibitor Standardization. Bovine trypsin and α-chymotrypsin were standardized according to the procedure of Estell et al. (1980) except that 4-methylumbelliferyl *p*-(*N*,*N*,*N*-trimethylammonio)cinnamate was substituted as the chymotrypsin burst titrant. The α-chymotrypsin was then used to standardize ovomucoid samples by titration at 10^{-7} M concentration. Several ovomucoid samples were then used to standardize elastase I and subtilisin by using the appropriate substrates at the higher concentrations.

Results

Ovomucoid Third Domain Sequence Comparison and Verification. The primary structures of isolated ovomucoid third domains from over 30 avian species have been determined in this laboratory by automated Edman degradation. Many of these sequences are listed in Kato et al. (1978). Comparisons of the sequences from Table IA provided eight pairs (Figure 1) such that each contains a different single amino acid substitution in the polypeptide chain. Additionally, if the first five residues are neglected (see Discussion), the set of pairwise comparisons can be extended to 12. See the single

Table I

A. Aligned Sequences of Ovomucoid Third Domains

Sequence Designator	Species Variant of Ovomucoid Third Domain	Sequence
		P_{18} P_{15} P_{10} P_5 P_1 P_1' P_5' P_{10}' P_{15}' P_{20}' P_{25}' P_{30}' P_{35}' P_{38}'
α	Chestnut Bellied Scaled Quail	F A A V S V D C S E Y P K P D C T L E Y R P L C G S D N K T Y A N K C N F C N A V V E S N G T L T L S H F G K C
β	Turkey	L A A V S V D C S E Y P K P A C T L E Y R P L C G S D N K T Y G N K C N F C N A V V E S N G T L T L S H F G K C
γ	Silver Pheasant	L A A V S V D C S E Y P K P A C T M E Y R P L C G S D N K T Y G N K C N F C N A V V E S N G T L T L S H F G K C
δ	Indian Peafowl	L A A V S V D C S E Y P K P A C T L E H R P L C G S D N K T Y G N K C N F C N A V V E S N G T L T L S H F G K C
ε	Guineafowl	L A A - - V D C S E Y P K P A C T M E Y R P L C G S D N K T Y D N K C N F C N A V V E S N G T L T L S H F G K C
ζ	Gambel's Quail	F A A V S V D C S E Y P K P A C T L E Y R P L C G S D N K T Y A N K C N F C N A V V E S N G T L T L S H F G K C
η	Mountain Quail	F A A V S V D C S E Y P K P A C T L E Y R P L C G S D S K T Y A N K C N F C N A V V E S N G T L T L S H F G K C
θ	Chukar	L A A - - V D C S E Y P K P A C T L E Y R P L C G S D S K T Y G N K C N F C N A V V E S N G T L T L S H F G K C
ι	Blue Scaled Quail	F A A V S V D C S E Y P K P D C T L E Y R P L C G S D N K T Y A S K C N F C N A V V E S N G T L T L S H F G K C
κ	Golden Pheasant	L A A V S V D C S E Y P K P A C T M E Y R P L C G S D N K T Y G M K C N F C N A V V E S N G T L T L N H F G K C
λ	Turkey (glycosylated)	L A A V S V D C S E Y P K P A C T L E Y R P L C G S D N K T Y G N K C N F C N A V V E S N G T L T L S H F G K C
μ	Silver Pheasant (glycosylated)	L A A V S V D C S E Y P K P A C T M E Y R P L C G S D N K T Y G N K C N F C N A V V E S N G T L T L S H F G K C
ν	Reeve's Pheasant	L A A V S V D C S E Y P K P A C T M E Y R P L C G S D N K T Y G M K C N F C N A V V E S N G T L T L S R F G E C
ξ	Ringnecked Pheasant	L A A V S V D C S E Y P K P A C T M E Y R P L C G S D N K T Y G N K C N F C N A V V E S N G T L T L N R F G Q C
ο	Duck	V A T - - V D C S G Y P K P A C T M E Y M P L C G S D N K T Y G N K C N F C N A V V D S N G T L T L S H F G E C
π	Goose	V A T - - V D C S D Y P K P A C T V E Y M P L C G S D N K T Y D N K C N F C N A V V D S N G T L T L S H F G K C
ρ	Rhea	F A T - - V D C S D H P K P V C S L E Y M P L C G S D S K T Y S N K C D F C N A V V E S N G T L T L S H F G K C

B. Position and Nearest Sequence Comparison

sequence designation	position of substitution and sequence compared
α	P_4:ξ or P_{15}':ι
β	P_1:γ or P_2':δ or P_{10}':θ or P_{14}':ζ or P_{27}':λ
γ	P_1:β or P_{14}':ε or P_{33}':κ or P_{27}':μ
δ	P_2':β
ε	P_{14}':γ
ζ	P_4:α or P_{14}':β or P_{10}':η
η	P_{10}':ζ or P_{14}':θ
θ	P_{10}':β or P_{14}':η
ι	P_{15}':α
κ	P_{33}':γ
λ	P_{27}':β
μ	P_{27}':γ
ν	P_{34}', P_{37}':γ or P_{33}', P_{37}':ξ
ξ	P_{33}, P_{34}, P_{37}':γ or P_{34}', P_{37}':κ or P_{33}', P_{37}':ν
ο	P_9, P_3, P_{25}' P_{37}':γ or P_9, P_1, P_{14}', P_{37}':π or P_9, P_3', P_{34}', P_{37}':ν
π	P_9, P_1, P_3', P_{27}':ε or P_9, P_1, P_{14}', P_{37}':ο
ρ	P_9, P_8, P_4, P_2, P_3', P_{10}', P_{14}', P_{18}':β or P_9, P_8, P_4, P_2, P_3', P_{14}', P_{18}':η, θ

replacements in Table III. Most of these changes can also be related one to another in single steps (Figure 1). Five other sequences have been included, but these contain more than one change to the nearest comparison sequence. From this set of 12 sequences, seven were redetermined, and all were checked against amino acid analysis to ensure accuracy.

Determination of Equilibrium Constants. Stoichiometries of enzyme and inhibitor were checked by titrations using an initial free enzyme concentration several orders of magnitude above the reciprocal of the association equilibrium constant. This was done to verify active, absolute concentrations and to check that the enzyme–inhibitor complex contained no residual small substrate activity which could be mistaken for complex dissociation. At the stoichiometric point, no residual activity was observed, and the operational normalities of all inhibitors and enzymes were self-consistent.

Equilibrium constants were determined at high dilution by the method of Green & Work (1953). The initial enzyme levels were generally within a factor of 10 of the inverse of K_a.

In the usual application of the Green & Work (1953) procedure, one starts with a single sample of enzyme and stepwise adds aliquots of the inhibitor solution. After an appropriate incubation time, a small test aliquot is withdrawn and placed in an assay medium (large dilution), and then the next aliquot of inhibitor is added. We believe that this procedure is incorrect for two reasons: (1) the incubation system, where equilibrium is reached, must be several times more concentrated than the system used for assay, and (2) the equilibrium mixture is perturbed prior to assay by a significant dilution. Both of these pitfalls were avoided here by making up separately each of the solutions defined in Figure 2 and then, after an appropriate incubation time, adding to them a small aliquot of concentrated substrate and commencing the assay.

The high dilution created several problems. First, extreme cleanliness had to be observed in order not to contaminate the assays. This was especially true for glassware and measuring devices which had previously been in contact with concentrated solutions of enzyme or inhibitor and then used for low concentrations. A more serious problem was that at the extremely low concentrations, the incubation vessels partially absorbed the samples. This effect has been observed recently by several authors (e.g., Johnson & Whatley, 1972; Freed & Ryan, 1980). The adsorption was a function of the surface area, charge, and hydrophilicity of the vessel wall. Several materials

Table II: Comparison of Third Domain Ovomucoid Enzyme Association Constants[a]

sequence designation	species variant of ovomucoid third domain	K_a (M^{-1}) chymotrypsin	elastase	subtilisin
α	chestnut bellied scaled quail	5.9×10^{11}	6.9×10^{8}	1.3×10^{6}
β	turkey	3.2×10^{11}	5.7×10^{10}	3.4×10^{10}
γ	silver pheasant	1.8×10^{11}	1.2×10^{10}	6.4×10^{10}
δ	Indian peafowl	1.3×10^{10}	1.3×10^{9}	2.7×10^{8}
ϵ	guinea fowl	4.0×10^{8}	1.4×10^{8}	4.2×10^{8}
ζ	Gambel's quail	7.2×10^{11}	1.1×10^{11}	3.7×10^{9}
η	mountain quail	5.2×10^{11}	9.5×10^{10}	3.6×10^{9}
θ	chukar	1.7×10^{11}	3.9×10^{10}	1.8×10^{10}
ι	blue scaled quail	4.0×10^{11}	3.7×10^{7}	2.0×10^{6}
κ	golden pheasant	1.2×10^{11}	1.2×10^{10}	1.9×10^{10}
λ	turkey (glycosylated)	3.5×10^{11}	4.4×10^{10}	2.7×10^{10}
μ	silver pheasant (glycosylated)	1.7×10^{11}	1.6×10^{10}	4.3×10^{10}
ν	Reeve's pheasant	3.4×10^{11}	1.1×10^{10}	2.2×10^{10}
ξ	ring-necked pheasant	7.3×10^{10}	1.2×10^{10}	7.4×10^{9}
o	duck	1.5×10^{10}	1.8×10^{11}	2.5×10^{10}
π	goose	1.8×10^{4}	2.0×10^{9}	1.0×10^{7}
ρ	rhea	8.9×10^{7}	1.8×10^{10}	7.6×10^{9}

[a] Data were obtained at 21 °C in buffer containing 0.2 M Tris, 0.02 M $CaCl_2$, and 0.005% (w/v) Triton X-100, pH 8.3.

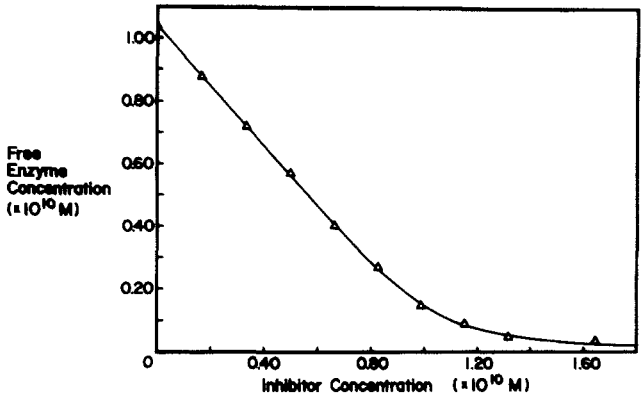

FIGURE 2: Titration curve of an initial constant amount of enzyme with incremental amounts of inhibitor (Green & Work, 1953). The remaining free enzyme is quantitated from turnover of a chromogenic substrate. Initial enzyme concentration employed is within a factor of 10 of the inverse K_a. Nonlinear least-squares generation of the best fit line through the data points yields the equilibrium constant K_a. These data are for chymotrypsin interacting with mountain quail ovomucoid third domain at 21 °C in buffer containing 0.2 M Tris, 0.02 M $CaCl_2$, and 0.005% (w/v) Triton X-100, pH 8.3.

were tried, with quartz and nitrocellulose being the best, but still unacceptable. It was found that addition of the nonionic detergent, Triton X-100, at 0.005% (w/v) concentration in the buffer prevented adsorption while not altering the kinetic properties of the enzymes. There was no alteration of enzyme–inhibitor parameters with Triton X-100 addition for weak inhibitors whose constants could be determined at higher concentrations where adsorption was not a problem.

The time required to reach equilibrium for the enzyme–inhibitor associations was calculated from the association rate constant. The most extreme assumptions were used, i.e., the time for stoichiometric amounts of enzyme and inhibitor to proceed through 20–30 half-lives of reaction. At high dilution, the required incubation time was on the order of 10–30 h. This length of time places a premium on negating any surface adsorption, since losses of enzyme activity greater than 10% significantly reduce accuracy of the determined K_a value. On the other hand, due to high dilution, autolysis of the enzymes is insignificant.

Titration curves of enzyme with inhibitor were fit (Figure 2) via a nonlinear least-squares analysis computer program to the theoretical equation. When these methods are used, the error in replicate experiments can be kept quite low. Specific discussion of the accuracy and the error in the equilibrium constant values is discussed at the end of this section. Generally, however, differences between two comparisons should be considered significant only if they are greater than a factor of 2 or the difference is supported by data obtained by using a different method.

Additionally, in cases where the equilibrium constants were weak ($<10^7$ M^{-1}), Green & Work (1953) titrations were carried out with large excesses of inhibitor and substrate concentrations below K_m values. These K_a determinations were further verified by substrate–inhibitor competition experiments according to the method of Estell et al. (1980).

The above experimental conditions may appear somewhat extreme. However, the Green & Work (1953) method has been frequently criticized for its potential pitfalls (e.g., Laskowski & Sealock, 1971; Bieth, 1974). The most obvious of these is attempting to interpret small differences from stoichiometric (straight line) behavior as a measure of K_a. Curvature of titration plots such as that shown in Figure 2 can arise from a number of causes unrelated to K_a. These causes include improper sample blanking, point scatter, inaccurate enzyme or inhibitor concentrations, residual activity of enzyme–inhibitor complex, contamination of the enzyme preparation by a small quantity of an enzyme which is not inhibited, insufficient incubation time to achieve enzyme–inhibitor equilibrium, or induced complex dissociation caused by substrate addition. Of all these points, the latter two are probably the most important. Incubation times should be calculated according to the second-order kinetic reaction; for complexes which have dissociation rate constants $k_d \gg 10^{-4}$ s^{-1}, substrate concentrations equal to or lower than $0.1K_m$ should be used. Reliable data are best obtained when these considerations are taken into account together with providing sufficient dilution such that 10–15% dissociation at the equivalence point is observed and all the data points are fit with low error to the theoretical equation.

Equilibrium Constant Comparisons. Single amino acid substitutions can result in striking modulations of the enzyme–inhibitor interaction. The results will be discussed in the sequential order of substitutions. Individual equilibrium constant values are given in Table II, and the ratio between the two compared equilibrium constants are listed in Table III. The positions of the substitution can be noted in Table IB and on the three-dimensional X-ray structure shown in Figure 3. The first comparison is that of an alanine–aspartic

Table III: Equilibrium Constant Factors Associated with Amino Acid Substitutions

sequences compared	position of substitution(s)	substitution $A_1 \rightarrow A_2$	ratio of equilibrium constants[a] (K_2/K_1) chymotrypsin	elastase	subtilisin
$\alpha \rightarrow \zeta$	P_4	Asp → Ala	1.2	160	2800
$\beta \rightarrow \gamma$	P_1'	Leu → Met	0.56	0.21	1.9
$\delta \rightarrow \beta$	P_2'	His → Tyr	25	44	130
$\zeta \rightarrow \eta$	P_{10}'	Asn → Ser	0.72	0.86	0.97
$\beta \rightarrow \theta$			0.53	0.68	0.53
$\beta \rightarrow \zeta$	P_{14}'	Gly → Ala	2.3	1.9	0.11
$\theta \rightarrow \eta$			3.0	2.4	0.20
$\epsilon \rightarrow \gamma$	P_{14}'	Asp → Gly	450	86	150
$\iota \rightarrow \alpha$	P_{15}'	Ser → Asn	1.5	19	0.65
$\beta \rightarrow \lambda$	P_{27}'	Asn → Asn[b]	0.94	1.3	0.67
$\gamma \rightarrow \mu$			1.1	0.77	0.79
$\gamma \rightarrow \kappa$	P_{33}'	Ser → Asn	0.67	1.0	0.30
$\gamma \rightarrow \nu$	P_{34}', P_{37}'	His → Arg, Lys → Glu	1.9	0.92	0.34
$\gamma \rightarrow \xi$	P_{33}', P_{34}', P_{37}'	Ser → Asn, His → Arg, Lys → Gln	0.41	1.0	0.12
$\gamma \rightarrow o$	P_9', P_3', P_{25}', P_{37}'	Glu → Gly, Arg → Met, Glu → Asp, Lys → Glu	0.08	15	0.39
$\beta \rightarrow \rho$	P_9, P_8, P_4, P_2', P_3', P_{10}', P_{14}', P_{18}'	Glu → Asp, Tyr → His, Ala → Val, Thr → Ser, Arg → Met, Asn → Ser, Gly → Ser, Asn → Asp	10^{-4}	0.39	0.22
$\epsilon \rightarrow \pi$	P_9, P_1', P_3', P_{25}'	Glu → Asp, Met → Val, Arg → Met, Glu → Asp	10^{-5}	14	0.02

[a] Data from Table II. K_2 is the equilibrium constant for A_2, and K_1 is the equilibrium constant for A_1. [b] Glycosylated.

acid change (Gambel's quail–chestnut bellied scaled quail, ζ–α) in sequence position P_4 (Schecter & Berger, 1967). The observed lowering of the inhibition constant by several orders of magnitude for elastase and subtilisin is consistent with the notion of extended binding sites for these enzymes (Robertus et al., 1972; Thompson & Blout, 1973). The P_4 position should contact the shallow hydrophobic pocket in elastase and the more extensive hydrophobic pocket in subtilisin. This would require removal of the negative charge on the aspartate site chain before burial in the enzyme pocket. The thermodynamic penalty of protonating the aspartic side chain at pH 8.3 is about 4 orders of magnitude, consistent with the observed weakened binding constants. The absence of an effect of the P_4 substitution on chymotrypsin is also consistent with its purported narrower binding site (Bauer, 1978). However, other workers suggest that chymotrypsin also recognizes the P_4 position (Wright, 1977). Since chymotrypsin contains a positively charged residue, Lys^{175}, in the P_4 contact region, it may be that the lack of change in K_a is a result of compensation between steric hindrance of the aspartate side chain and of possible ion pair stabilization.

The comparison of the primary recognition residue, P_1, is between leucine and methionine (turkey–silver pheasant, β–γ). The amino acid change also exerts a differential effect on specificity. Both elastase and chymotrypsin are bound more strongly by a leucine reactive site while subtilisin is bound more tightly by methionine. This is not unexpected since the specific subtilisin inhibitor, streptomyces subtilisin inhibitor, which exhibits a high homology both structurally and sequentially in regions of enzyme contact to ovomucoid third domains, also contains a methionine reactive site (Ikenaka et al., 1974).

The binding pocket in elastase is purported to be highly restrictive (Shotton & Watson, 1970). Obviously, the binding pocket is not too highly restrictive since the leucine reactive site still exhibits strong inhibitory properties and since small peptide substrates with leucine and large or branched hydrophobic residues have good k_{cat}/K_m ratios (Del Mar et al., 1979; Thomson & Kapadia, 1979). This problem will be further discussed (Papamokos et al., 1982). The differences between leucine and methionine are quite small because these two

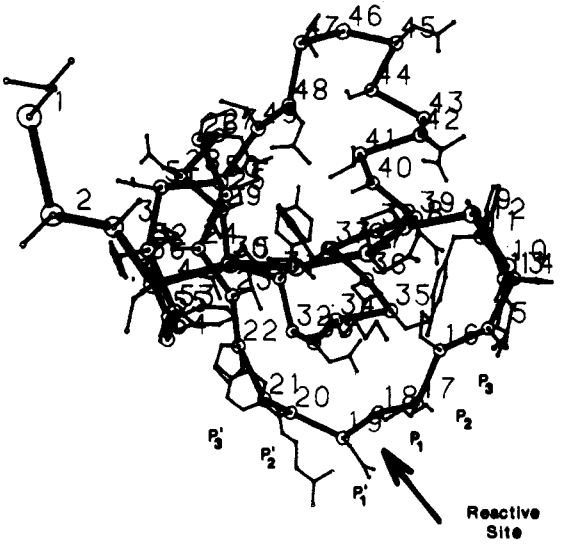

FIGURE 3: Three-dimensional structure of Japanese quail ovomucoid third domain (Weber et al., 1981). The reactive site is labeled by an arrow and by the Schecter & Berger (1967) notation. Arabic numerals are the residue positions beginning from the N terminus of the 56-residue domain (see Table IA).

amino acid side chains are so similar. However, truly dramatic differences can arise from P_1 substitutions by other amino acids. As an example, even though the comparision is flawed, note how dramatically goose ovomucoid third domain with valine at P_1 differs from the leucine and methionine inhibitors.

At position P_2', a normally conserved tyrosine change to a histidine (turkey–Indian peafowl, β–δ) results in a decreased binding constant for all three enzymes. Similarly, introduction of a charged aspartic acid residue from a glycine (guinea fowl–silver pheasant, ϵ–γ) in P_{14}' also decreases the binding constants for each of these enzymes by several orders of magnitude. Again, a less hydrophobic residue can exert a large effect. At position P_{14}' a second comparison can be made. The alanine to glycine change (Gambel's quail–turkey, ζ–β, and mountain quail–chukar, η–θ) decreases the binding constants for elastase and chymotrypsin by a factor of 2 but increases the constant for subtilisin by a factor of 10, suggesting a steric interaction in subtilisin. This is a striking example of differential specificity and binding constant modulation by a single residue sequentially far from the primary recognition site. This clearly demonstrates that all residue side chains which come in contact with the enzyme potentially can exert large effects on binding. Even small conservative changes like an alanine to a glycine replacement may have large effects.

The position immediately adjacent, P_{15}', compares an asparagine to a serine change (chestnut bellied scaled quail–blue scaled quail, α–ι). Here, only elastase is affected, and the effect is to decrease the binding constant by a factor of 19. In the C terminus, a P_{33}' serine is changed to an asparagine (silver pheasant–golden pheasant, γ–κ). The observed inhibition constants show only subtilisin to be slightly affected.

The final two comparisons are at position P_{10}' and P_{27}', respectively, an asparagine to serine change (Gambel's quail–mountain quail, ζ–η, and turkey–chukar, β–θ) and, the most radical, the attachment of a bulky carbohydrate group to an asparagine (turkey, β–λ, and silver pheasant, γ–μ). Neither of these replacements has an effect on binding. Although both are surface residues, their locations preclude contact with any of the enzymes tested. This suggests that only the surface residue changes in positions which contact the enzyme need be considered as potential modulators.

The association equilibrium constants for five additional ovomucoid third domains have been determined. However, these domains exhibit more than one amino acid substitution compared to its nearest sequence, and unfortunately it appears that in each case more than one change is responsible for differences in K_a. In two of these comparisons, the changes take place in the last six residues of the C terminus (silver pheasant–Reeve's pheasant, γ–ν, P_{34}', P_{37}'; silver pheasant–ring-necked pheasant, γ–ξ, P_{33}', P_{34}', P_{37}'). The observed effects are only slight, suggesting this region of the inhibitor may make only a weak contact with the enzyme. The third comparison (duck–silver pheasant, o–γ) contains four substitutions. Two positions are expected to have little effect (P_{25}' and P_{37}'); the remaining two, P_9 and P_3', both may contact the enzyme, with P_3' expected to make the strongest contact. The equilibrium constants show a strong differential effect. It is interesting to note that the P_3' residue may contact the enzyme near a threonine (residue 61) in chymotrypsin and an arginine (residue 61) in elastase. The change of the P_3' arginine to a methionine (along with the other substitutions) increases the binding constant for elastase by a factor of 16 and decreases the binding for chymotrypsin by a factor of 12. However, the other changes may also be contributing to the observed modulation.

The fourth comparison is between rhea and turkey, with eight changes including P_2 and P_3'. The observed binding constants are slightly reduced for elastase and subtilisin and strongly reduced for chymotrypsin. Both P_2 and P_3' are expected to exert strong effects.

The last comparsion is between goose and guinea fowl (π–ϵ). Substitutions exist at P_9, P_1, P_3', and P_{25}', with the P_1 and P_3' positions expected to be most significant. It is striking that the goose ovomucoid third domain is such a weak inhibitor of chymotrypsin and subtilisin yet still strong for elastase. Part of this effect may be due to the methionine in P_3', based on the above data. This would leave a huge effect by the P_1 residue change of a leucine to valine. This reaffirms the great importance of P_1 residues in determining inhibitor binding strength. For subtilisin, the goose inhibitor is sufficiently weak that it exhibits temporary inhibition. (The temporary inhibition was also seen for the blue scaled quail and chestnut bellied scaled quail samples with subtilisin.)

The tendency of weak subtilisin inhibitors being temporary can probably be explained by the relatively low specificity of subtilisin which could render more bonds other than the reactive site susceptible to hydrolysis and by the rapid formation by subtilisin of modified inhibitors (W. Ardelt and M. Laskowski, Jr., unpublished results; see also the introduction). It is interesting to note that fragmentary results indicate that carbohydrate-containing ovomucoid third domains are less susceptible to temporary inhibition than the carbohydrate free third domains, indicating that bonds on the opposite side of the inhibitor molecule from the reactive site (see Figure 3) are likely substrates in the temporary inhibition phenomenon.

An additional interesting point can be seen by inspection of Table II. It is well-known that many, but not all, protein proteinase inhibitors which inhibit chymotrypsin also inhibit subtilisin. In most of these cases, it has been shown that chymotrypsin is more strongly inhibited than subtilisin. Most of the entries in Table II clearly follow this generalization. However, there are two well-publicized exceptions, streptomyces subtilisin inhibitor (S-SI) (Inouye et al., 1979) and penguin ovomucoid (Osuga et al., 1974). To this list we now add the two bottom entries in Table II—goose ovomucoid third domain and rhea ovomucoid third domain. The molecular explanation for goose third domain is clearly P_1 Val, but for rhea, as stated above, it cannot as yet be unambiguously determined.

Rate Constant Determination. Accuracy of the equilibrium constant values was verified by determining the association, k_{on}, and dissociation, k_d, rate constants (Table IV); their ratio defines the equilibrium constant. Association rate constants were determined directly by mixing equimolar amounts of enzyme and inhibitor at low concentration in the presence of a 10-fold lower than K_m concentration of substrate. The relation of inverse free enzyme levels with time gave a linear second-order plot through at least three half-lives of reaction. Dissociation rate constants were determined by first forming enzyme–inhibitor complex, inducing dissociation by adding a large excess of burst substrate for chymotrypsin (see Methods), methyl-His[57] chymotrypsin (Ryan & Feeney, 1975) for elastase, or a 10-fold lower than the inverse K_a concentration of complex to saturating amounts of substrate for subtilisin, and quantitating either free enzyme or inhibitor with time. The appropriate first-order plots were linear through three half-lives.

A comparison of the data shows the directly determined and calculated equilibrium constant values to be consistent within the error of the method (Table V). It should be noted that

Table IV: Rate Constants for Ovomucoid Third Domain-Enzyme Interaction[a]

ovomucoid	chymotrypsin[b]		elastase[b]		subtilisin[c]	
	k_{on} (M^{-1} s^{-1})	k_d (s^{-1})	k_{on} (M^{-1} s^{-1})	k_d (s^{-1})	k_{on} (M^{-1} s^{-1})	k_d (s^{-1})
silver pheasant	1.3×10^7	4.5×10^{-5}	1.7×10^6	1.2×10^{-4}	1.5×10^6	2×10^{-5}
turkey	1.2×10^7	2.5×10^{-5}	1.1×10^6	3.4×10^{-5}	1.1×10^6	3×10^{-5}
turkey with carbohydrate	0.9×10^7	2.7×10^{-5}	1.1×10^6	3.2×10^{-5}	9.0×10^5	3×10^{-5}
Gambel's quail	1.3×10^7	1.0×10^{-5}	1.2×10^6	1.3×10^{-5}	1.4×10^6	2×10^{-4}

[a] Data were obtained at 21 °C in buffer containing 0.2 M Tris, 0.02 M $CaCl_2$, and 0.005% (w/v) Triton X-100, pH 8.3. [b] Dissociation rate constants for chymotrypsin and elastase are for k_{off} only. [c] Dissociation rate constants for subtilisin are for the sum of k_{off} and k_{off}^* (W. Ardelt, unpublished results). Association rate constant k_a is for virgin inhibitor only.

Table V: Comparison of Calculated and Directly Determined Equilibrium Constants[a]

ovomucoid third domain	chymotrypsin		elastase		subtilisin	
	$K_{kinetic}$ (M^{-1})	K_{direct} (M^{-1})	$K_{kinetic}$ (M^{-1})	K_{direct} (M^{-1})	$K_{kinetic}$ (M^{-1})	K_{direct} (M^{-1})
silver pheasant	2.9×10^{11}	1.8×10^{11}	1.4×10^{10}	1.2×10^{10}	7.5×10^{10}	6.4×10^{10}
turkey (unglycosylated)	4.8×10^{11}	3.2×10^{11}	3.2×10^{10}	5.7×10^{10}	3.7×10^{10}	3.4×10^{10}
turkey (glycosylated)	3.3×10^{11}	3.5×10^{11}	3.4×10^{10}	4.4×10^{10}	3.0×10^{10}	2.7×10^{10}
Gambel's quail	1.3×10^{12}	7.2×10^{11}	0.9×10^{11}	1.1×10^{11}	7.0×10^9	3.7×10^9

[a] $K_{kinetic} = k_{on}/k_d$, data from Table IV; K_{direct} data from Table II.

Table VI: Comparison of Equilibrium Constant Ratios Obtained by Competition Experiments[a] and by Direct Determination

sequence comparison (1 → 2)	enzyme	ratio of equilibrium constants K_2/K_1	
		competition	from Table III
$\beta \rightarrow \lambda$	α-chymotrypsin	1.0	0.94
$\beta \rightarrow \gamma$	α-chymotrypsin	0.59	0.56
$\beta \rightarrow \lambda$	elastase I	1.0	1.3
$\beta \rightarrow \gamma$	elastase I	0.08	0.21
$\gamma \rightarrow \mu$	subtilisin Carlsberg	0.83	0.79
$\beta \rightarrow \gamma$	subtilisin Carlsberg	1.9	1.9

[a] Enzyme and inhibitors were incubated at 21 °C in 0.02 M Tris buffer at pH 8.3.

when equilibrium constants are affected by substitutions, it is the dissociation rate constant only which is changed. Particularly striking was the measurement for an inhibitor with a charged residue such as an aspartate in position P_4 or P_{14}' (data not shown). This resulted in k_d increases of several orders of magnitude. A similar effect on k_d has been observed for trypsin–soybean trypsin inhibitor interactions (Laskowski et al., 1976) and may imply similar explanations for mechanism.

Equilibrium Competition. Direct competition experiments further substantiated the accuracy of the relative equilibrium constants given in Table III. In these experiments, precisely equimolar amounts of one inhibitor and of enzyme were mixed and allowed to form complex. To this complex a precisely equimolar amount of the second inhibitor was added and allowed to compete with the bound inhibitor for the enzyme. It can be shown that the ratio of the amount of the two complexes at equilibrium is equal to the square root of the ratio of the respective equilibrium constants (K_a).

Specifically, after being mixed, aliquots containing microgram quantities of protein were withdrawn at various times and subjected to gel exclusion, high-pressure liquid chromatography. Monitoring was done at 206 nm to increase sensitivity. The chromatography traces at initial and equilibrium times are shown in Figure 4. Figure 5 is a plot of the time course of the reaction. From these figures, it is seen that the same equilibrium position is attained from both directions; that

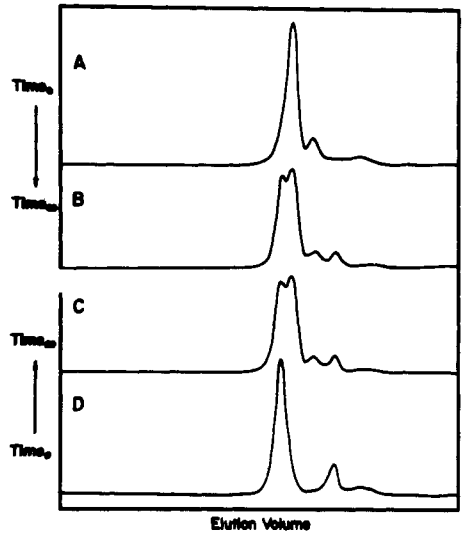

FIGURE 4: High-pressure liquid chromatography elution profiles monitored at 206 nm for equilibrium competition experiments. (A) Free glycosylated inhibitor (silver pheasant third domain) and enzyme nonglycosylated inhibitor complex (α-chymotrypsin and turkey third domain) at initial mixing. (B) Mixture A after equilibrium is achieved. The two major peaks are glycosylated and nonglycosylated inhibitor–enzyme complex, respectively. (C) Mixture D after equilibrium is achieved. (D) Free nonglycosylated inhibitor and enzyme glycosylated inhibitor complex at initial mixing.

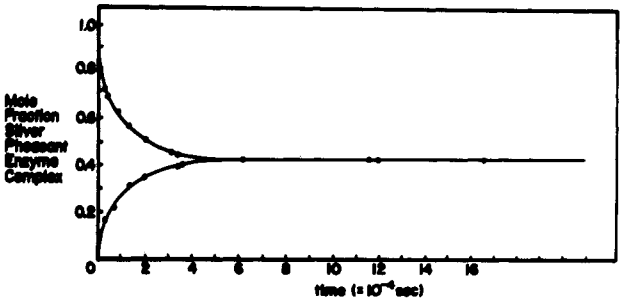

FIGURE 5: Time course of equilibrium competition for turkey (nonglycosylated) and silver pheasant (glycosylated) ovomucoid third domain with α-chymotrypsin. The upper curve is generated by preforming silver pheasant third domain α-chymotrypsin complex and adding the turkey inhibitor to it. The distribution of products with time is analyzed as shown in Figure 4. The lower curve is generated in an analogous manner except the turkey–α-chymotrypsin complex is first preformed with the silver pheasant inhibitor added to it.

is, the equilibrium position is independent of the order of inhibitor addition. A comparison of the ratios of equilibrium constants obtained by this technique with those of Table III is given in Table VI. It is seen that the agreement is generally satisfactory.

In these experiments, carbohydrate attached to P_{27}' served as a neutral marker by greatly increasing the apparent Stokes radius. This provided for the facile separation of the four components (nonglycosylated inhibitor, glycosylated inhibitor, complex of enzyme and nonglycosylated inhibitor, and complex of enzyme and glycosylated inhibitor). The presence of carbohydrate was shown to be nonperturbing since it has a negligible effect upon K_a (Table III, $\beta \rightarrow \lambda$ and $\gamma \rightarrow \mu$) and upon the kinetic constants k_{on} and k_d (Table IV).

This technique has two disadvantages: (1) If one of the two association constants is quite weak, k_d will be large, and the 15-min chromatographic run may significantly disturb the original equilibrium composition. This should not be the case for k_d values less than 10^{-4} s^{-1} (half-life greater than 100 min); thus, weakly associating pairs may not be probed by this technique. (2) If the ratio of the equilibrium constants is greater than 100 (or less than 0.01), one of the complexes is present at equilibrium at only 9% of the total amount of the two complexes. Due to imperfections in the base line of the chromatograms (Figure 4), it is very difficult to quantitate the small peak corresponding to the weak inhibitor complex. Hence, unreliable results for the ratio of the equilibrium constants of such a mixture are obtained. Outside of these limitations, the technique appears sound and accurate.

Error Analysis. In order to examine Table III in detail, it is very useful to be aware of the range of experimental errors in order not to overinterpret the meaning of the entries. Five methods of assessing such an error are available:

(1) Each primary data set such as Figure 2 yields not only the value of log K_a but also a probable error in log K_a. The average probable error in the data we have utilized is ±0.1 (range 0.02–0.30). This is a measure of the goodness of fit of the actual points to the best curve through the data.

(2) All of the K_a values listed in Table II and then used to calculate the ratios entered in Table III were determined at least in duplicate and frequently in triplicate. From the reproducibility an average error of log K_a ± 0.1 (range 0.02–0.2) was calculated. These two statistics are both measures of random error in K_a determination. Note that the entries in Table III are ratios of K_a's, and therefore the error of the determination may be propagated in the ratio.

(3) The K_a values in 12 cases were also determined from kinetic measurements (see above, Table IV) in order to check for systematic errors. The error in the log K_a ratio calculated from this set is also ±0.1 (range −0.25 to 0.28), suggesting that systematic errors are probably smaller than random errors. It is also possible that $K_{kinetic}$ is more accurate than K_{direct}.

(4) In six cases, the values of the ratio of K_a's were rechecked by direct competition experiments. The observed values were consistent with the K_a's determined by the other methods (see Table VI). The error in log K_a is ±0.1 (range −0.4 to +0.03).

(5) For three single amino acid replacements (positions P_{10}', P_{14}', and P_{27}'), it was possible to determine the effect of the same replacement by two different systems. These were internally consistent and the *ratios* of equilibrium constants show an error of log K_1/K_2 = ±0.1 (range 0.07–0.26) (for the data see Table III).

In the above discussion all measurements were treated as equivalent, thus generating a sufficient statistical base. However, when individual measurements are considered, it becomes apparent that (a) the largest (and also possibly the smallest) K_a's have the largest error and (b) the results for subtilisin are generally more scattered than for chymotrypsin and elastase. This is understandable since in the subtilisin system, virgin to modified inhibitor conversion is an important reaction, while it is too slow to influence the results for chymotrypsin and elastase (W. Ardelt and M. Laskowski, Jr., unpublished results).

On the basis of the discussion given above, we have adopted a rather conservative interpretation. We state that if the ratios listed in Table III lie in the range 0.5–2.0 (Δ log ratio ±0.3) there is no experimentally measurable effect if the data are based on K_a determinations only. However, we interpret some ratios in this range as significantly different from unity if the same answer is obtained both by direct K_a determinations and by competition experiments or by kinetically determined K_a's.

Discussion

This paper differs from the previous literature on sequence to reactivity relationships of proteinase inhibitors in several important respects. First, most of the compared inhibitors differ from one another by only a single amino acid replacement, thus eliminating complex (and often faulty) reasoning required to assign the change in reactivity to one of several replacements. Second, the inhibitory domains which are compared were subjected to very rigorous characterization to assure that they were pure and that their structure corresponded to the listed sequence. Third, the equilibrium constants for association with enzymes and the corresponding rate constants were determined as accurately as possible and subjected to error analysis. Fourth, in the set of comparisons described here, changes at positions other than P_1 and P_1' were emphasized. Previous studies, which tracked single amino acid replacements, were primarily based on variants obtained by semisynthetic replacements of P_1 and P_1' positions (Sealock & Laskowski, 1969; Kowalski et al., 1974; Kowalski & Laskowski, 1976; Odani & Ikenaka, 1978; Odani & Ono, 1980; Wenzel & Tschesche, 1981). While these papers were based primarily on qualitative or only semiqualitative comparisons, they have clearly established the primary role of the P_1 residue and only marginal importance of P_1' residue in determination of the strength and specificity of association. In this paper we show that aside from P_1 many other positions are important both to the strength and to specificity of inhibitors.

It is the intention of our laboratory to generate (both by isolation and by semisynthesis) a large set of ovomucoid domains each differing from another member of the set by a single amino acid replacement and to determine for each member of this set the thermodynamic and kinetic data for interaction with various enzymes. With a sufficiently large data base, we expect to be able to predict these properties on the basis of their amino acid sequence alone for all other ovomucoid domains which are not members of the data set. We call this goal the sequence → reactivity algorithm (Laskowski, 1980; Laskowski et al., 1981). At present the data set described in this paper is very small, and thus the predictive power is very limited. However, we can already rationalize why chicken ovomucoid third domain is not an efficient inhibitor of any of the three enzymes utilized in this study (K_a < 10^7 M^{-1}). Chicken ovomucoid third domain differs from turkey ovomucoid third domain (Table IA, β) at three positions: at P_4 it has Asp instead of Ala, at P_1 it has Ala instead of Leu, and at P_2' it has Asp instead of Tyr. The presence of Asp at positions P_4 and P_2' should greatly weaken the binding of subtilisin and of elastase, while both the P_1 Ala and

the P_2' Asp are very unfavorable for binding of chymotrypsin.

The data in Table III (the equilibrium constant ratios) provide values for the $\Delta\Delta G°$ of transfer of side chains from the surface of the inhibitor to the enzyme–inhibitor interface. As this interface is mainly hydrophobic, it should be expected that the $\Delta\Delta G°$ values should roughly follow the $\Delta G°$ for transfer of side chains from the aqueous to nonaqueous environment (Nozaki & Tanford, 1971). While this trend is generally followed, several exceptions can be seen (positions P_4 for chymotrypsin and P_{14}' for subtilisin). It is clear that in order to account for such deviations one must include effects of steric hindrance, possible hydrogen bonds, and salt bridges. Such information can be obtained from the three-dimensional structure of enzyme–inhibitor complexes determined by X-ray crystallography or inferred from the equilibrium constant ratios themselves, provided that a large enough number of side chains are investigated at each position. At present, the three-dimensional structures of enzyme inhibitor complexes directly relevant to this work were not yet determined, although several such determinations are underway. To substitute for it, models of all the complexes described here were generated (Papamokos et al., 1982) on a computer from the known three-dimensional structures of the free enzymes and from the three-dimensional structure of Japanese quail ovomucoid third domain (Weber et al., 1981). These models indicate that where a residue replacement affects K_a the side chain at that position makes contact with the enzyme and the positions where replacements do not affect K_a do not make contact. This strong statement should be qualified slightly since the models are not equivalent to direct determination of three-dimensional structure and since in some cases (because of experimental accuracy of K_a) it is not clear whether a significant effect is observed or not. Furthermore, all large effects on K_a can be readily rationalized on the basis of interactions suggested by the models. As more X-ray crystallographic data becomes available, it seems likely that the values in Table III can be generated by energy calculations and compared to experimental data. We think that this system may be particularly worthwhile for energy calculations as (1) the effects observed here are probably as highly localized as possible as opposed to global effects, such as an overall equilibrium constant for denaturation of an entire protein, and (2) the structure of both reactants (enzyme and inhibitor) and of the product (the enzyme–inhibitor complex) can be determined by X-ray crystallography, thus avoiding such poorly defined states as "denatured protein". It is clear that it might be helpful to determine as well the $\Delta H°$ values for the associations listed in Table II because then one could compare the $\Delta\Delta H°$ and the $\Delta\Delta S°$ values for each side chain replacement. Similarly, we plan to extend the analysis to the determination of the various rate constants characterizing the enzyme–inhibitor interaction and to the values of $K_{hyd}°$, the equilibrium constant for the reactive site hydrolysis. However, such extensions appear at present to be even more labor intensive than the determination of the K_a values reported here.

In making the first 12 comparisons listed in Table III, we say that the compared domains differ by only one amino acid replacement. This is, however, strictly true for only eight comparisons (see Table IA). In an additional four comparisons (and in many other distant comparisons), we neglect the differences in the first five residues in the NH_2-terminal region. The differences are between Phe^1 and Leu^1 in some cases and the deletion of the dipeptide Val^4-Ser^5 vs. its presence in some others. It should be noted that the neglected residues are not properly a part of the third domain anyway but constitute part of the connecting peptide between the second and third domain. The first residue of the true third domain is Val^7 (Stein et al., 1980). The justification for this neglect is based on three lines of reasoning: (1) The X-ray structure of Japanese quail third domain shows that residues 1–6 are not a part of the integral structure of the domain and furthermore that these residues are quite distant from the reactive site (Weber et al., 1981). (2) The computer confrontation models (Papamokos et al., 1982) show that residues 1–6 are far from making contact with any of the enzymes studied here. (3) Different proteinases cut the connecting peptide between second and third domains at different positions (Kato et al., 1978). Yet, all such "third domains" are active as inhibitors, and in qualitative tests no difference in inhibitory activity or specificity has been detected.

The neglect of the first five residues need not remain a "skeleton in our closet". Experiments are now planned, which will allow us to put this neglect to an experimental test, but we feel quite confident that no effect will be observed.

The choice of ovomucoid third domains to investigate the contribution which specific amino acid residues make toward enzyme–inhibitor interaction was based on several characteristics of the third domains. Many of these are not present in other sets of inhibitors. The third domains can be easily isolated, readily characterized, and rapidly sequenced. They are small, compact molecules, and they are very soluble and highly stable and, we assume, very rigid. The three-dimensional structure of one of them is known (Weber et al., 1981), and several others will soon be determined. In their interaction with proteinases they follow the "standard mechanism" (Laskowski & Kato, 1980), thus facilitating the determination of the parameters for the enzyme inhibitor interaction. However, the major advantage is the hypervariability of amino acid residues in the critical enzyme inhibitor contact region (Kato et al., 1978; Bogard et al., 1980). Thus, many of the available variants had their single replacements in the contact region which also led to significant differences in K_a values.

This study allowed us to notice another interesting evolutionary trend. Quite frequently, the fixation of one mutation which greatly weakens the interaction with some enzyme is rapidly followed by fixation of another mutation which weakens this interaction even further. This is shown by the following example. Chestnut bellied scaled quail third domain (α) differs from closely related Gambel's quail third domain (ζ) only by P_4 Asp rather than by P_4 Ala (which is typical for most phasianoid birds). The P_4 Ala → Asp fixation weakens the binding of elastase and subtilisin. This mutation is followed in blue scaled quail (ι) by a P_{15}' Asn → Ser replacement (very uncommon). This replacement further weakens the interaction, with elastase making the inhibitor almost ineffective, while the inhibition of chymotrypsin remains strong and unaffected by either of the fixed mutations. Another somewhat more difficult example of the same phenomenon is the three changes between turkey ovomucoid and chicken ovomucoid third domain. As stated above, chicken ovomucoid is not an effective inhibitor of any of the three enzymes studied here while turkey is highly effective for all three. For each enzyme, chicken has *two* highly deleterious changes when compared to turkey, whose sequence appears to be near that of the ancestral phasianoid bird. The phenomenon appears to be more general than the examples given here, but a more detailed discussion requires a construction of a complete phylogenetic tree.

We hope that the ultimate product of this and subsequent work from our laboratory will be a detailed sequence to reactivity algorithm. Such an algorithm can be constructed by

the methods described here only if the parameters describing the interaction of the various inhibitor variants with the test enzymes are highly accurate. Considerable effort in this paper has gone toward improving methods of measurement of K_a and assessing their accuracy. In spite of considerable improvement in accuracy over what is now reported in the literature, our accuracy is only marginally sufficient and further improvement will probably be needed.

Conclusions

(1) Replacement of surface residues in positions remote from enzyme–inhibitor contact area by other side chains has no effect upon K_a. (2) Replacement of surface residues in the enzyme–inhibitor contact area shows significant effects upon K_a even when the residue positions are sequentially remote from the reactive site. (3) Such replacements occasionally exert strong differential effects weakening interaction with one enzyme and strengthening with another. (4) Even with the present small data base, the ultimate chance of success for a sequence to reactivity algorithm appears promising. As an example, the present data can be used to explain the low inhibitory activity of chicken ovomucoid third domain with the three enzymes studied here.

Acknowledgments

We thank Dr. S. Blumberg and Dr. C. Largman for their generous gifts of substrates, Dr. I. Kato for sequence determinations, Dr. W. Finkenstadt for computer programming, and Dr. W. Ardelt for many helpful discussions.

References

Bauer, C. A. (1978) *Biochemistry 17*, 375–380.
Bieth, J. (1974) *Bayer-Symp. 5*, 463–469.
Bieth, J., Spiess, B., & Wermuth, C. G. (1974) *Biochem. Med. 11*, 350–357.
Bogard, W. C., Jr., Kato, I., & Laskowski, M., Jr. (1980) *J. Biol. Chem. 255*, 6569–6574.
Chase, T., Jr., & Shaw, E. (1967) *Biochem. Biophys. Res. Commun. 29*, 508–514.
Del Mar, E. G., Largman, C., Brodrick, J. W., & Geokas, M. C. (1979) *Anal. Biochem. 99*, 316–320.
Empie, M. W., & Laskowski, M., Jr. (1981) *Fed. Proc., Fed. Am. Soc. Exp. Biol. 40*, 1601.
Estell, D. A., Wilson, K. A., & Laskowski, M., Jr. (1980) *Biochemistry 19*, 131–137.
Feeney, R. E. (1971) *Proc. Int. Res. Conf. Proteinase Inhibitors, 1st*, 189–195.
Freed, R. C., & Ryan, D. S. (1980) *Biochim. Biophys. Acta 624*, 562–572.
Green, N. M., & Work, E. (1953) *Biochem. J. 54*, 347–352.
Hoy, T. G., Ferdinand, W., & Harrison, P. M. (1974) *Int. J. Pept. Protein Res. 6*, 121–140.
Ikenaka, T., Odani, S., Sakai, M., Nabeshima, Y., Sato, S., & Murao, S. (1974) *J. Biochem. (Tokyo) 76*, 1191–1209.
Inouye, K., Tonomura, B., & Hiromi, K. (1979) *J. Biochem. (Tokyo) 85*, 601–607.
Jameson, G. W., Roberts, D. V., Adams, R. W., Kyle, S. A., & Elmore, D. T. (1973) *Biochem. J. 131*, 107–117.
Johnson, P., & Whatley, T. L. (1972) *Biochim. Biophys. Acta 276*, 323–327.
Kato, I., Schrode, J., Wilson, K. A., & Laskowski, M., Jr. (1976) *Protides Biol. Fluids 23*, 235–243.
Kato, I., Kohr, W. J., & Laskowski, M., Jr. (1978) *Proc. FEBS Meet. 47*, 197–206.
Kowalski, D., & Laskowski, M., Jr. (1976) *Biochemistry 15*, 1309–1315.
Kowalski, D., Leary, T. R., McKee, R. E., Sealock, R. W., Wang, D., & Laskowski, M., Jr. (1974) *Bayer-Symp. 5*, 311–324.
Laskowski, M., Jr. (1980) *Biochem. Pharmacol. 29*, 2089–2094.
Laskowski, M., Jr., & Sealock, R. W. (1971) *Enzymes, 3rd Ed. 3*, 375–473.
Laskowski, M., Jr., & Kato, I. (1980) *Annu. Rev. Biochem. 49*, 593–626.
Laskowski, M., Jr., Finkenstadt, W. R., Mattis, J. A., & McKee, R. E. (1976) *Protides Biol. Fluids 23*, 85–91.
Laskowski, M., Jr., Empie, M. W., Kato, I., Kohr, W. J., Ardelt, W., Bogard, W. C., Jr., Weber, E., Papamokos, E., Bode, W., & Huber, R. (1981) *32nd Mosbach Colloquium* (Eggerer, H., & Huber, R., Eds.) pp 136–152, Springer-Verlag, Berlin.
Lineweaver, H., & Murray, C. W. (1947) *J. Biol. Chem. 171*, 565–581.
Luthy, J. A., Praissman, M., Finkenstadt, W. R., & Laskowski, M., Jr. (1973) *J. Biol. Chem. 248*, 1760–1771.
Nozaki, Y., & Tanford, C. (1971) *J. Biol. Chem. 246*, 2211–2217.
Odani, S., & Ikenaka, T. (1978) *J. Biochem. (Tokyo) 84*, 1–9.
Odani, S., & Ono, T. (1980) *J. Biochem. (Tokyo) 88*, 1555–1558.
Osuga, D. T., Bigler, J. C., Uy, R. L., Sjoberg, L., & Feeney, R. E. (1974) *Comp. Biochem. Physiol. B 48B*, 519–533.
Papamokos, E., Weber, E., Bode, W., Huber, R., Empie, M. W., Kato, I., & Laskowski, M., Jr. (1982) *J. Mol. Biol.* (in press).
Pozsgay, M., Gaspar, R., Bajusz, S., & Elodi, P. (1979) *Eur. J. Biochem. 95*, 115–119.
Quast, U., Engel, J., Steffan, E., Tschesche, H., & Kupfer, S. (1978) *Eur. J. Biochem. 86*, 353–360.
Rhodes, M. B., Bennet, N., & Feeney, R. E. (1960) *J. Biol. Chem. 235*, 1686–1693.
Robertus, J. D., Kraut, J., Alden, R. A., & Birktoff, J. J. (1972) *Biochemistry 11*, 4293–4304.
Ryan, D. S., & Feeney, R. E. (1975) *J. Biol. Chem. 250*, 843–847.
Schecter, I., & Berger, A. (1967) *Biochem. Biophys. Res. Commun. 27*, 157–162.
Sealock, R. W., & Laskowski, M., Jr. (1969) *Biochemistry 8*, 3703–3710.
Shinar, S., & Gertler, A. (1979) *Int. J. Pept. Protein Res. 13*, 218–222.
Shotton, D. M., & Watson, H. C. (1970) *Nature (London) 225*, 811–817.
Stein, J. P., Catterall, J. F., Kristo, P., Means, A. R., & O'Malley, B. W. (1980) *Cell (Cambridge, Mass.) 21*, 681–687.
Thompson, R. C., & Blout, E. R. (1973) *Biochemistry 12*, 66–71.
Thomson, A., & Kapadia, S. B. (1979) *Eur. J. Biochem. 102*, 111–116.
Wang, C.-C., & Shaw, E. (1972) *Arch. Biochem. Biophys. 150*, 259–268.
Weber, E., Papamokos, E., Bode, W., Huber, R., Kato, I., & Laskowski, M., Jr. (1981) *J. Mol. Biol. 149*, 109–123.
Wenzel, H. R., & Tschesche, H. (1981) *Angew. Chem., Int. Ed. Engl. 20*, 295–296.
Wright, H. T. (1977) *Eur. J. Biochem. 73*, 567–578.