Interscaffolding Additivity. Association of P_1 Variants of Eglin c and of Turkey Ovomucoid Third Domain with Serine Proteinases†

M. A. Qasim,‡ Philip J. Ganz,§ Charles W. Saunders,§ Katherine S. Bateman,|| Michael N. G. James,*,|| and Michael Laskowski, Jr.*,‡

Department of Chemistry, Purdue University, West Lafayette, Indiana 47907, Miami Valley Laboratories, Procter & Gamble Company, Cincinnati, Ohio 45253, and Medical Research Council of Canada Group in Protein Structure and Function, Department of Biochemistry, University of Alberta, Edmonton, Alberta T6G 2H7, Canada

Received August 19, 1996; Revised Manuscript Received December 9, 1996⊗

ABSTRACT: Standard mechanism protein inhibitors of serine proteinases share a common mechanism of interaction with their cognate enzymes. The P_1 residue of the inhibitor interacts with the enzyme in a substrate-like manner. Its side chain becomes imbedded in the S_1 cavity of the enzyme. The nature of P_1, the primary specificity residue, greatly affects the strength and specificity of the enzyme inhibitor association. In canonical inhibitors, residues $P_4-P_2'(P_3')$, where P_1-P_1' is the reactive site, share a common main chain conformation that does not change on complex formation. The remainder of the inhibitor's structure, the scaffolding, is not always common. Instead, there are at least 20 inhibitor families, each with a different scaffolding. In this paper, we ask whether the differences in standard free energy of association of enzyme−inhibitor complexes upon P_1 mutations are independent of the nature of the scaffolding. We have already reported on 25 P_1 variants of turkey ovomucoid third domain, a member of the Kazal inhibitor family, interacting with six different serine proteinases. Here, we report on seven different P_1 variants of eglin c, a potato I family member, interacting with the same six serine proteinases under the same conditions. The differences in standard free energy on P_1 mutations in the eglin c system agree very well, when P_1 Pro is omitted. Complete agreement indicates that these P_1 residues are interscaffolding additive. This is consistent with the superimposition of the high-resolution structures of eglin c and of turkey ovomucoid third domain with chymotrypsin. In both cases, the P_1 Leu side chain is similarly oriented in almost indistinguishable specificity pockets of the enzyme.

Association of standard mechanism (Laskowski & Kato, 1980), canonical (Bode & Huber, 1992) protein inhibitors of serine proteinases with their cognate enzymes involves about a dozen of the residues of the inhibitor (Figure 1) making contact with the enzyme. Standard mechanism inhibitors are ones where the enzyme−inhibitor complex can reversibly dissociate into two different sets of products. One set consists of the enzyme and the original inhibitor, with all of its peptide bonds intact. Such an inhibitor is called the virgin inhibitor. The other set consists of the enzyme and the inhibitor with one of its peptide bonds, called the reactive site peptide bond, hydrolyzed. Such an inhibitor is called the modified inhibitor. It is an assertion of the standard mechanism that the virgin and modified inhibitors are both comparably thermodynamically efficient as inhibitors, although (Ardelt & Laskowski, 1985) they may dramatically differ in their kinetic efficiency. In the Schecter and Berger (1967) notation, which is now widely used by students of proteinase substrates and inhibitors, the reactive site peptide bond connects residues P_1 and P_1'. Other residues are numbered sequentially by their distance from the reactive site (Figure 1). A large number of free inhibitors and an even larger number of enzyme−inhibitor complexes had their three-dimensional structures determined by X-ray crystallography, NMR, or both. A general conclusion has emerged that residues $P_4-P_2'(P_3')$ exhibit the same Ramachandran angles in many free inhibitors studied. Upon complex formation, these Ramachandran angles do not appreciably change. Bode and Huber (1992) named inhibitors with this conformation of $P_4-P_2'(P_3')$ canonical inhibitors. It is now commonly believed that being a standard mechanism inhibitor is necessary and sufficient for being canonical.

Upon complex formation, the P_1 residue of the inhibitor becomes imbedded in the S_1 cavity of the enzyme. In most natural trypsin inhibitors, the P_1 residue is Arg or Lys. In chymotrypsin inhibitors, it is an aromatic amino acid or Leu or Met, and in porcine pancreatic elastase inhibitors, it is Leu or Ala. Therefore, the P_1 residue in substrates and in protein inhibitors is often referred to as the primary specificity residue and the S_1 cavity as the primary specificity cavity. Sealock and Laskowski (1969) devised the first method for replacing amino acid residues in proteins. They replaced P_1 Arg^{63} in soybean trypsin inhibitor (Kunitz) by P_1 Lys^{63} and found that it was still a powerful trypsin inhibitor. Leary and Laskowski (1973) switched the P_1 Arg^{63} residue to P_1 Trp^{63} and as expected found it to be a powerful chymotrypsin inhibitor. Since the advent of more elaborate semisynthesis, total synthesis, and, most of all, site specific mutagenesis, many replacements were made at P_1 in many laboratories. They were frequently successful in switching the specificity of the inhibitor in a desired manner.

† Supported at Purdue by NIH Grant GM10831 and at Alberta by grants from the Medical Research Council of Canada and The Alberta Heritage Foundation.

* Corresponding authors.
‡ Purdue University.
§ Procter & Gamble.
|| University of Alberta.

⊗ Abstract published in *Advance ACS Abstracts,* February 1, 1997.

S0006-2960(96)02087-9 CCC: $14.00

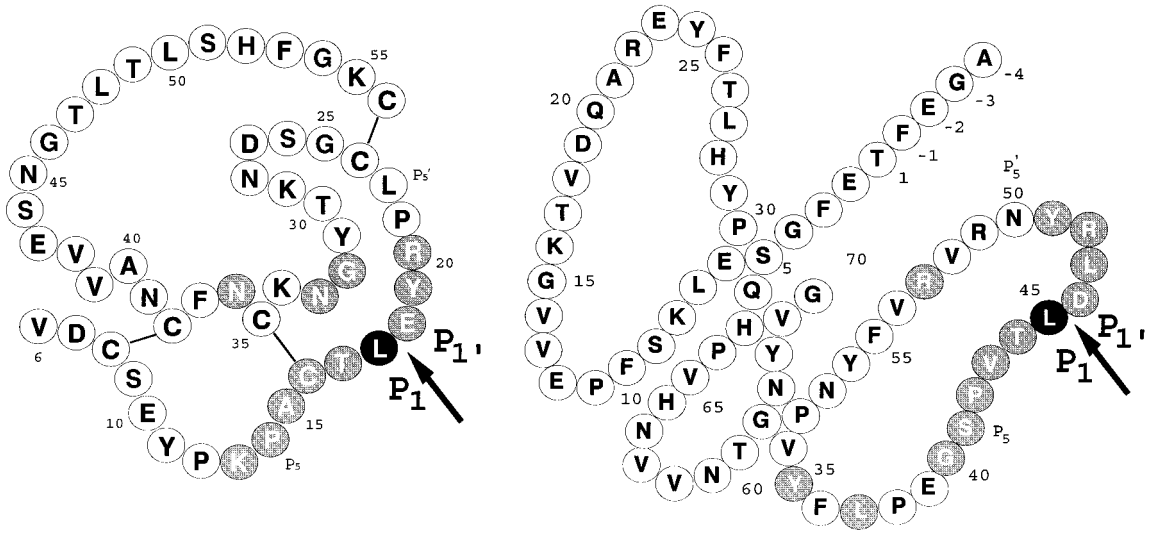


FIGURE 1: Covalent structures of turkey ovomucoid third domain, OMTKY3 (Laskowski et al., 1987) (left-hand side), and of eglin c (Seemüller et al., 1980) (right-hand side). The recombinant OMTKY3 variants whose K_a values are utilized in this paper start at residue 6. The four NH_2 terminal residues in eglin c marked −4 to −1 are a result of recombinant DNA operations (see Materials and Methods). The arrows indicate the reactive site peptide bonds, which serve as origins of the Schecter and Berger (1967) P_n−P_n' notation. In each inhibitor, the consensus[2] set of residues that come in contact with the cognate enzyme are darkened. The P_1 Leu in both inhibitors (Leu^{18} and Leu^{45}, respectively) are in black. The − indicates disulfide bridges in OMTKY3. There are no disulfide bridges in eglin c.

A single-domain, standard mechanism, canonical protein inhibitor can inhibit only one enzyme molecule at a time. However, just as one kind of enzyme can be inhibited by many different kinds of inhibitors, so can one kind of inhibitor inhibit many different kinds of enzymes. Structural studies and especially chemical studies (Bidlingmeyer et al., 1972; Ardelt & Laskowski, 1985) show that all the inhibited enzymes utilize the same reactive site peptide bond and the same P_1 residue is imbedded in the different S_1 cavities of the various cognate enzymes.

If the P_1 residue is mutated to Gly, the positive interaction between the S_1 cavity and the P_1 side chain disappears as Gly has no side chain. Mutation to a very awkwardly shaped or to a potentially charged side chain may actually make the P_1 side chain−S_1 cavity interaction locally deleterious.

Turkey ovomucoid third domain (OMTKY3)[1] (Figure 1) is a standard mechanism, canonical protein inhibitor of serine proteinases. We have prepared 25 P_1 mutants (all 20 coded, 5 noncoded) of this protein. We studied (Lu et al., 1997) their binding to six serine proteinases at pH 8.3 and 21 ± 2 °C. Among the six enzymes, one was SGPB. It turns out that, of the 25 variants we studied, three (Pro, Asp, and Glu) are locally deleterious. We obtained high-resolution structures of many of the 25 variants in complex with SGPB (Huang et al., 1995; Huang, 1995; K. Huang, K. S. Bateman, W. Lu, S. Anderson, M. Laskowski, Jr., and M. N. G. James, unpublished). Among the completed structures are P_1 Gly^{18} (neutral) and the three locally deleterious ones (P_1 Pro^{18}, P_1 Asp^{18}, and P_1 Glu^{18}). In each case, the P_1 side chain binds to the S_1 cavity and the P_1 O atom binds to the oxyanion hole. The available evidence is consistent with the view that the standard mechanism canonical inhibitors act as molecular vises; they deliver the P_1 side chain of the inhibitor to the S_1 cavity of the enzyme whether (as it generally is) the interaction is favorable, or even when it is neutral or locally deleterious.

As discussed above, Lu et al. (1997) acquired a set of 150 (6 × 25) enzyme−inhibitor equilibrium constants. The obvious question to be addressed is how general are the data. It is clear that the association equilibrium constants do not *directly* apply to other inhibitors. It is well known that bovine pancreatic trypsin inhibitor (Kunitz), BPTI, with a Lys^{15} at P_1 is one of the most powerful natural inhibitors of bovine β-trypsin with a K_a of $>10^{13}$ M^{-1} (Vincent & Lazdunski, 1972; Finkenstadt et al., 1974). In sharp contrast, chicken ovomucoid *first* domain with P_1 Lys^{24} is ineffective as an inhibitor of β-trypsin (Kato et al., 1987). Clearly, residues other than P_1 exert huge cumulative effects on K_a. The P_1 K_a data for X^{18} OMTKY3 can still be general if the *relative* K_a values apply to other inhibitors. The constancy of relative equilibrium constants is totally equivalent to the additivity of free energy contributions due to the P_1 residue.

The concept of additivity of residue free energy contributions arose many times in many laboratories that were interested in comparing properties of multiple mutants of their protein to those of the same protein's individual single mutants. We used it for analysis of K_a's of natural variants of avian ovomucoid third domains (Laskowski et al., 1983). Carter et al. (1984) used it for site specific mutants of Tyr

[1] Abbreviations: CHYM, bovine chymotrypsin Aα; PPE, porcine pancreatic elastase; CARL, subtilisin Carlsberg; SGPA and SGPB, *Streptomyces griseus* proteinases A and B, respectively; HLE, human leukocyte elastase; rmsd, root mean square deviation; PDB, Protein Data Bank; EDTA, ethylenediaminetetraacetic acid; Tris, tris(hydroxymethyl)aminomethane; MES, 2-(*N*-morpholino)ethanesulfonic acid. Avian ovomucoid third domains are abbreviated OMXXX3, where the first two letters indicate ovomucoid, the next three show the species of bird, and the final digit signifies the domain. TKY is turkey, *Melleagris gallopovo*, and MNQ is Montezuma or harlequin quail, *Cyrtonyx montezumae*. BPTI is bovine pancreatic trypsin inhibitor. Residues in OMTKY3 and in eglin c are numbered both sequentially and by Schecter and Berger (1967) with the P_n and P_n' notation from their reactive sites (see Figure 1). When the distinction between enzyme and inhibitor residues is required, the inhibitor residues are preceded by I.

[2] The consensus is among the complexes of eglin c with subtilisin Carlsberg, CARL, subtilisin BPN′, thermitase, mesentericopeptidase, and chymotrypsin (Bode et al., 1987; McPhalen & James, 1988; Gros et al., 1989; Heinz et al., 1991; Dauter & Betzel, 1991; Frigerio et al., 1992).

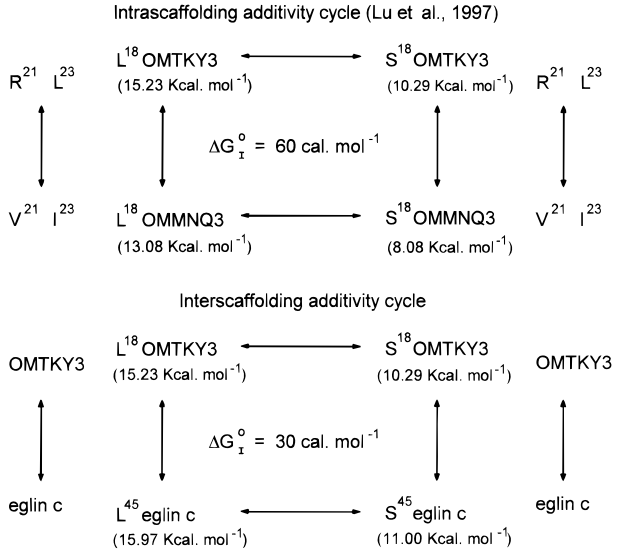


FIGURE 2: Examples of intrascaffolding and interscaffolding additivity cycles involving changes in P_1 residues. The abbreviations OMTKY3 and OMMNQ3 stand for turkey ovomucoid third domain and Montezuma quail ovomucoid third domain, respectively. The difference between the $\Delta\Delta G_a°$(Leu P_1 Ser) values in the first and second lines is the intrascaffolding $\Delta G_I°$. The difference between the third and fourth lines gives interscaffolding $\Delta G_I°$. The enzyme is CHYM.

tRNA synthetase and Ackers and Smith (1985) for comparison of mutant human hemoglobins. In a famous review, Wells (1990) defines additivity as the absence of the $\Delta G_I°$ term [interaction energy of Carter et al. (1984)] in

$$\Delta\Delta G_{(X,Y)}° = \Delta\Delta G_{(X)}° + \Delta\Delta G_{(Y)}° + \Delta G_I° \qquad (1)$$

$\Delta\Delta G_{(X,Y)}°$ is the difference in $\Delta G°$ values between the combined mutant X,Y and the wild type, and $\Delta\Delta G_{(X)}°$ and $\Delta\Delta G_{(Y)}°$ are the differences between the mutants X and Y and the wild type, respectively. The X mutation and Y mutation clearly must be made at different positions, but they need not necessarily be single mutations. For our purposes, mutation X will be a single change of the P_1 residue. However, Y can be a single change or several changes at positions other than P_1 as long as the same changes are made in both the combined X,Y and the Y mutant. From the extensive work of natural variants of avian ovomucoid third domains, we had available data for the construction of many additivity cycles. Figure 2 (top) may serve as an example. It was found that, within the available group of avian ovomucoids, the P_1 substitutions were highly, although not perfectly, additive [Figure 7 in Lu et al. (1997)]. It is therefore clear that the data should be helpful in designing multiple variants of OMTKY3 as strong and highly specific inhibitors of serine proteinases. Unfortunately, data were not conveniently available to extend the additivity tests to other members of the Kazal inhibitor family such as pancreatic secretory trypsin inhibitors from various vertebrates, ovoinhibitor domains, and even avian ovomucoid first and second domains.

However, not all standard mechanism canonical protein inhibitors of serine proteinase are homologous as might have been expected. Instead, they consist of many families (Laskowski & Sealock, 1971; Laskowski & Kato, 1980; Laskowski, 1986). As more inhibitors are discovered, the number of families is growing. At the current count, there

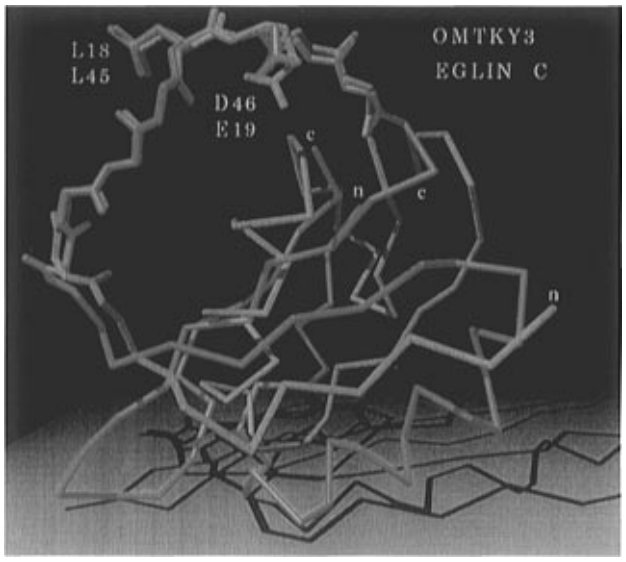


FIGURE 3: Superimposition of eglin c (blue) and OMTKY3 (green) as these inhibitors are bound to chymotrypsin. The backbone atoms (N, CA, C, and O) are shown for residues P_6-P_3' (I13−I21, OMTKY3; I40−I48, eglin c). Side chain atoms are included for P_1-P_1' residues. The remainders of the two inhibitors are depicted as a CA trace. This figure was made with RASTER3D (Bacon & Anderson, 1988).

are 20 different families. The relationship between families is best illustrated in Figure 3, where the combining regions of OMTKY3 and of eglin c and their P_1 Leu residues are superimposed. In contrast to the expected excellent superimposition, the global three-dimensional structures are entirely different. There is no recognizable sequence homology. We refer to the nonsuperimposable global structures as the scaffolding. If additivity comparisons are made among inhibitors that share the entire global conformation, we talk about intrascaffolding additivity. If, however, additivity comparisons are made between two different scaffoldings, we talk about interscaffolding additivity (Figure 2, bottom). The objective of this paper is to carry out a test of interscaffolding additivity.

In order to make interscaffolding comparisons, we needed a moderately large set of P_1 variants for an inhibitor other than a Kazal family member. The Procter and Gamble group had a set of seven P_1 variants of eglin c and agreed to cooperate with us. Eglin c (Figure 1) is a member of the potato I family (Laskowski & Kato, 1980). Its sequence and scaffolding (Figures 1 and 3) are strikingly different from that of OMTKY3, as eglin c has no disulfide bridges and is stabilized solely by noncovalent interactions. On the other hand, the main chain conformations of the combining loops of the two inhibitors superimpose superbly (Figure 3). The sequences in this region are also similar (Figure 1). Both inhibitors have P_1 Leu (Leu^{18} and Leu^{45}, respectively). In this paper, we measure K_a values for the available eglin c variants with six serine proteinases studied by Lu et al. (1997). We compare the results to the OMTKY3 results and show that the changes at P_1 are highly additive except for P_1 Pro. An explanation for the lack of additivity of this residue is proposed.

Both eglin c and OMTKY3 have been subject to extensive structural studies. In particular, high-resolution structures of eglin c (Frigerio et al., 1992) and of OMTKY3 (Fujinaga et al., 1987) with the same enzyme, bovine chymotrypsin Aα, were reported. As the thermodynamic comparisons are

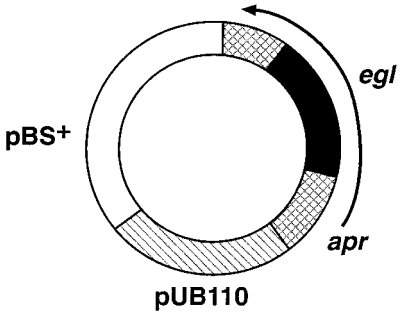


FIGURE 4: To produce eglin variants, we used *Bacillus subtilis* strains carrying a plasmid with four functional parts. One, the plasmid contained most of pBS+ (open section) which enabled it to replicate, confer ampicillin resistance, and be used as a template for oligonucleotide-directed mutagenesis in *Escherichia coli*. Two, the plasmid contained pUB110 (hatched section) which enabled it to replicate and confer kanamycin resistance in *B. subtilis*. Three, the plasmid contained the *B. amyloliquefaciens* subtilisin gene (*apr*) regulatory signals and signal sequence coding region (double-hatched section). Four, the plasmid contained a synthetic eglin gene (*egl*) (closed section).

focused on the binding of P_1 side chains to S_1 cavities, we carried out a superimposition of these regions in both structures. The finding of near identity of the conformation of the S_1 pocket of chymotrypsin and of the conformation of the P_1 Leu side chain in that pocket greatly raises the confidence in the validity of the thermodynamic conclusion that, for Leu and several other side chains, this is an additive system.

MATERIALS AND METHODS

Expression of Eglin c Variants. A synthetic eglin c gene was constructed which was identical to the *Eco*RI to *Bam*HI synthetic eglin c gene fragment described by Rink et al. (1984) except that we omitted the ATG adjacent to the *Eco*RI site.

The expression plasmid used (see Figure 4) was based on a pUB110/pBS$^+$ shuttle vector containing the *Bacillus amyloliquefaciens* subtilisin gene (Vasantha et al., 1984). The plasmid derived from a plasmid containing (1) 3.2 kb of pBS+, from its *Hin*dIII site to its *Xba*I site, (2) a 4.2 kb fragment of pUB110 from its *Xba*I site to its *Bam*HI site, (3) a 33 bp fragment of a synthetic human parathyroid gene (Saunders et al., 1991) from its *Bgl*II site to its *Hin*dIII site, and (4) a 1.8 kb segment consisting of the subtilisin gene, flanked by *Hin*dIII sites. *Eco*RI and *Bam*HI sites were removed from this plasmid by digestion with those enzymes, treatment with T4 polymerase in the presence of four deoxyribonucleotide triphosphates, and ligation.

The subtilisin gene was modified to contain an *Eco*RI site following the second codon of the prosequence, as identified by Vasantha et al. (1984), and a *Bam*HI site following the stop codon. These changes were made using the oligonucleotides 5′-GCC CAG GCG GCA GGG GAA TTC AAA TCA AAC GGG GAA-3′ and 5′-GCG GCA GCT CAG TAA GGA TCC AAC ATA AAA AAC CGG C-3′. The synthetic eglin c gene was cloned into the expression plasmid between the *Eco*RI and *Bam*HI sites.

Because of the nature of the gene fusion, each eglin c additionally contains four extra amino acid residues at its amino terminus: Ala and Gly from the prosequence of subtilisin and Glu and Phe from the *Eco*RI site at the fusion (see Figure 1).

Site-Directed Mutagenesis. Site-directed mutagenesis was carried out using the Kunkel method (Sambrook et al., 1989). The following oligonucleotides were used to create the indicated changes at the P_1 position of eglin c: Ser, 5′-TCT CCT GTT ACT AGT GAC CTG CGT TAC AAC CG-3′; Ile, 5′-TCT CCT GTT ACT ATC GAT CTG CGT TAC AAC-3′; Gly, 5′-TCT CCT GTT ACT GGA GAT CTG CGT TAC AAC-3′; Asp, 5′-TCT CCT GTT ACT GAT GAC CTG CGT TAC AAC CGT-3′; Glu, 5′-TCT CCT GTT ACT GAA GAT CTG CGT TAC AAC-3′; and Pro, 5′-TCT CCT GTT ACT CCG GAC CTG CGT TAC AAC CGT-3′. All mutations were confirmed by DNA sequencing using Applied Biosystems' Ready Reaction Dye Deoxy Terminator Cycle Sequencing Kit and an Applied Biosystems 373A DNA Sequencer.

Purification of Eglin c Variants. Eglin c variants were harvested from plasmid-bearing derivatives of *Bacillus subtilis* strain PG632 (Saunders et al., 1991). The culture (200 mL) was incubated overnight at 37 °C in 2 X YT (Sambrook et al., 1989) supplemented with 50 μg/mL kanamycin. The cell pellet was removed by centrifugation, and the pH of the supernatant was brought to 3 with the addition of 1 N HCI. The insoluble material was pelleted again by centrifugation at 22000*g*, and the supernatant was filtered through a 0.2 μm membrane. The sample was dialyzed overnight against 20 mM sodium acetate at pH 4. The sample was then purified on an S Sepharose column in the same buffer. Elution was performed with a 0 to 0.6 M sodium chloride gradient. The sample was concentrated by evaporation and dialyzed against 1 mM MES and 0.5 mM EDTA at pH 6.2. The identity of samples was confirmed by ion spray mass spectrometry.

K_a Determination. Determinations of equilibrium constants were carried out by an extensively modified procedure of Green and Work (1953), as described by Empie and Laskowski (1982) with later modifications by Park (1985) and Wynn (1990). CHYM, CARL, and HLE were purchased from Worthington Biochemical Co., Sigma Chemical Co., and Elastin Products Co., respectively. PPE was a gift from the late Dr. M. Laskowski, Sr. SGPA and SGPB were purified from pronase (Sigma) in our laboratory. The results on all commercial enzymes were compared to those on enzymes obtained from expert (see Acknowledgment) donors. The commercial samples were used only when the results agreed. For most enzymes we studied, the procedure allows for the determination of K_a values over 10 orders of magnitude (K_a from about 10^3 to 10^{13} M^{-1}) with an accuracy of ±20%. The substrate hydrolysis data were collected on a Hewlett-Packard HP8450A diode array spectrophotometer at 21 ± 2 °C and pH 8.30 in 0.1 M Tris/HCI containing 0.005% Triton X-100 and 0.02 M $CaCl_2$. The residual enzyme concentrations in the incubation mixture were measured using appropriate tri- or tetrapeptide *p*-nitroanilide chromogenic substrates.

Superimpositions. Atomic coordinates were obtained from the PDB (Bernstein et al., 1977) (CHYM and eglin c, lacb; CHYM and OMTKY3, 1cho). The crystal forms used to determine the structures of the two complexes were non-isomorphous. Therefore, the crystal packing interactions are expected to be different. InsightII was used to superimpose the core chymotrypsin residues from the active site and S_1 pocket. Specifically, atoms from residues 41−43, 55−59, 101−104, 189−196, 213−217, and 225−228 were chosen

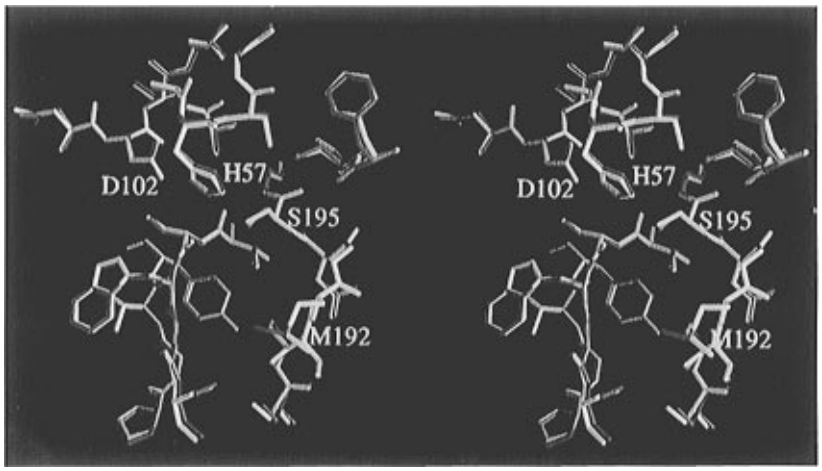


FIGURE 5: Stereoscopic view of the chymotrypsin residues which were used to calculate the superimposition of CHYM and eglin c (red) and CHYM and OMTKY3 (yellow). This figure was generated with GRASP (Nicholls et al., 1991).

Table 1: Equilibrium Constants (M^{-1}) for the Seven P_1 Variants of Eglin c Interacting with Six Serine Proteinases[a]

P_1 variants	CHYM	PPE	CARL	SGPA	SGPB	HLE
Gly (0)	2.2×10^7	1.0×10^6	1.9×10^{10}	3.1×10^{10}	2.6×10^{10}	9.5×10^8
	6.6×10^6	*9.0×10^8*	*6.4×10^8*	*4.5×10^7*	*1.2×10^7*	*2.2×10^7*
	3.3	**0.0011**	**30**	**690**	**2200**	**43**
Ser (2)	1.4×10^8	5.0×10^6	1.2×10^{11}	1.9×10^{11}	2.9×10^{11}	8.1×10^9
	4.2×10^7	*8.6×10^8*	*2.4×10^9*	*1.6×10^8*	*5.0×10^7*	*3.3×10^7*
	3.3	**0.0058**	**50**	**1200**	**5800**	**250**
Pro (3)	5.4×10^6	9.0×10^4	9.8×10^6	3.3×10^7	3.9×10^7	2.9×10^8
	6.7×10^4	*5.9×10^5*	*7.9×10^4*	*4.6×10^4*	*3.6×10^4*	*1.9×10^5*
	81	**0.15**	**124**	**720**	**1600**	**1500**
Leu (4)	6.7×10^{11}	1.0×10^8	8.9×10^{11}	$>10^{13}$	$>10^{13}$	1.8×10^{11}
	1.9×10^{11}	*4.2×10^{10}*	*3.4×10^{10}*	*3.0×10^{11}*	*5.6×10^{10}*	*6.1×10^9*
	3.5	**0.0024**	**26**			**30**
Ile (4)	4.8×10^8	4.1×10^7	5.8×10^9	3.3×10^{10}	4.7×10^{10}	2.7×10^{11}
	9.9×10^7	*5.9×10^9*	*1.4×10^8*	*1.5×10^8*	*2.9×10^7*	*2.1×10^{10}*
	4.9	**0.0069**	**41**	**220**	**1600**	**13**
Asp (4)	6.2×10^6	<500	5.2×10^8	1.1×10^9	3.7×10^9	3.7×10^5
	1.0×10^6	*7.3×10^4*	*1.6×10^7*	*5.1×10^6*	*3.9×10^6*	*1.6×10^4*
	6.2		**33**	**220**	**950**	**23**
Glu (5)	9.0×10^6	<500	1.6×10^{10}	2.4×10^9	1.7×10^9	3.7×10^5
	2.5×10^6	*8.4×10^4*	*8.6×10^8*	*7.9×10^6*	*2.3×10^6*	*4.1×10^4*
	3.6		**19**	**300**	**740**	**9**.0

[a] In each case, the data in italics are for a corresponding variant of OMTKY3 (Lu et al., 1997) and the number in bold is the eglin c/OMTKY3 ratio. All the eglin c and OMTKY3 data were determined directly by the same method at pH 8.3 and $t = 21 \pm 2$ °C. Numbers in parentheses are the non-hydrogen atoms in each side chain. These data were converted to $-\Delta G_a°(P_1 = X)$, and the remainder of the paper uses only those values.

(Figure 5). Atom CE of Met^{192} was not included in the calculation as this atom points in the opposite direction in the two complexes. Root mean square deviations (rmsds) were calculated with the difres program from the CCP4 program suite.

RESULTS

K_a Values. The seven P_1 eglin c variants were expressed and purified as described. Their masses determined by ion spray mass spectrometry were in excellent agreement with those expected from their amino acid sequences (Figure 1, modified at P_1 as appropriate).

Table 1 lists K_a values for the interaction of each of these seven variants with each of the six serine proteinases employed by Bigler et al. (1993) and Lu et al. (1997) in their extensive study of the interaction of 25 P_1 variants of OMTKY3. The 25 variant OMTKY3 set contains all seven P_1 residues found in the eglin c set. To further improve the comparability of the eglin c and of the OMTKY3 sets, the K_a values for both inhibitors were determined by the same method, in the same laboratory, under common conditions (pH, temperature, and ionic environment).

In the OMTKY3 work (Lu et al., 1997), K_a values were determined for all 150 (25×6) combinations. All the values fell within the $500-10^{13}$ M^{-1} dynamic range of measurement. Surprisingly, this rather broad range proved insufficient for some of the eglin c variants. SGPA and SGPB are much more strongly inhibited by eglin c variants than by corresponding OMTKY3 variants. The K_a value for the eglin c P_1 Leu variant interacting with either SGPA or SGPB is greater than 10^{13} M^{-1}. PPE is much more weakly inhibited by eglin c variants than by corresponding variants of OMTKY3. The K_a values for the P_1 Asp and Glu eglin c variants are lower than 500 M^{-1}.

Inspection of Table 1 shows that inhibition by corresponding eglin c and OMTKY3 variants is not the same. This was expected as eglin c does not have the same contact residues at many positions other than P_1. Eglin c is better for five of the six enzymes: SGPB > SGPA > HLE > CARL > CHYM. It is much weaker for PPE. However,

Table 2: $\Delta G_I°$ for the Seven P_1 Variants of OMTKY3 and Eglin c with Six Different Serine Proteinases

	CHYM	PPE	CARL	SGPA	SGPB	HLE
G → S	0	970	310	320	580	1020
G → P	1870	2890	840	20	−400	2090
G → L	30	450	−70	ND	ND	−220
G → I	220	1070	200	−670	−170	−710
G → D	360	ND[a]	50	−680	−480	−370
G → E	50	ND	−270	−480	−630	−920
S → P	1870	1920	530	−300	−980	1070
S → L	30	−520	−380	ND	ND	−1240
S → I	220	100	−110	−990	−750	−1730
S → D	360	ND	−250	−1000	−1060	−1390
S → E	50	ND	−580	−800	−1210	−1940
P → L	−1840	−2440	−910	ND	ND	−2300
P → I	−1650	−1810	−640	−690	240	−2800
P → D	−1500	ND	−790	−700	−10	−2460
P → E	−1820	ND	−1110	510	−220	−3010
L → I	190	630	270	ND	ND	−490
L → D	330	ND	130	ND	ND	−140
L → E	10	ND	−200	ND	ND	−700
I → D	140	ND	−140	−10	−310	340
I → E	−180	ND	−470	190	−460	−210
D → E	−320	ND	−330	200	−150	−550

[a] ND, not determined.

in each column, the ratios of corresponding variants are approximately the same (with some strong exceptions dealt with below). This is analyzed quantitatively below. It may be useful to point out that, if the system showed perfect interscaffolding additivity, the ratios in each column would be identical. Another way to word the same conclusion is to say that the relative equilibrium constants within each variant set would be identical.

Interscaffolding Additivity. In order to serve as input for additivity comparisons, the K_a data of Table 1 were converted into changes in standard free energy of association by the use of the relation $\Delta G_a°\ (P_1 = X) = -RT \ln K_a(P_1 = X)$. From these values, increments arising from mutating P_1 residues were calculated. Consider the two horizontal lines of the bottom part of Figure 2. For these we can write

$$\Delta\Delta G_a°(\text{Leu } P_1 \text{ Ser})_{\text{OMTKY3}} = \Delta G_a°(P_1 = \text{Ser})_{\text{OMTKY3}} - \Delta G_a°(P_1 = \text{Leu})_{\text{OMTKY3}} \quad (2)$$

and

$$\Delta\Delta G_a°(\text{Leu } P_1 \text{ Ser})_{\text{eglin c}} = \Delta G_a°(P_1 = \text{Ser})_{\text{eglin c}} - \Delta G_a°(P_1 = \text{Leu})_{\text{eglin c}} \quad (3)$$

It follows from Figure 2 and eqs 1−3 that

$$\Delta\Delta G°(\text{Leu } P_1 \text{ Ser})_{\text{OMTKY3}} = \Delta\Delta G°(\text{Leu } P_1 \text{ Ser})_{\text{eglin c}} + \Delta G_I° \quad (4)$$

where $\Delta G_I°$ is the free energy of interaction. For the specific case of Leu → Ser mutation, $\Delta G_I°$ can be calculated for each of the six enzymes. However, as we are dealing here with seven different P_1 residues, we could calculate $\Delta G_I°$ for 21 [(7 × 6)/2] combinations, each for six enzymes, provided all K_a values were available. The lack of four eglin c K_a values reduces the number of available $\Delta G_I°$ values to 103 from the hoped for 126. These are listed in Table 2. If the system were to be strictly additive, there would be no standard free energy of interaction and all of the entries in Table 2 would be zero. It is clear that except for G → S for chymotrypsin all other values are not zero.

Table 3: (A) Intrascaffolding Additivity Involving P_1 Substitutions in Ovomucoid Third Domains (Lu et al., 1997) and (B) Interscaffolding Additivity Involving P_1 Substitutions in Eglin c and Turkey Ovomucoid Third Domain

(A) intrascaffolding			(B) interscaffolding				
			with Pro			without Pro	
$\lvert\Delta G_I°\rvert$	no.	%	$\lvert\Delta G_I°\rvert$	no.	%	no.	%
0−400	44	67.7	0−400	47	45.6	41	57.8
410−800	16	24.6	410−800	24	23.3	18	25.4
810−1200	4	6.3	810−1200	14	13.6	9	12.7
>1200	1	1.5	>1200	18	17.5	3	4.2
total	65	100	total	103	100	71	100

Experimental Errors Also Produce $\Delta G_I°$ Terms. As is seen in both parts of Figure 2, every additivity cycle involves comparison of four different free energy terms. Measurement of these terms involves errors. We estimate that the errors in our K_a determinations are about ±20% at the σ level. This corresponds to about ±100 cal/mol as the σ level error in $\Delta G_a°$ values. The error in $\Delta G_I°$ is therefore the propagated error of four such measurements and about ±200 cal/mol. In Bigler et al. (1993) and Lu et al. (1997), we decided that the 2 σ level is better for examination of deviations from additivity. We therefore regard $\Delta G_I°$ values in Table 2 between −400 and +400 cal/mol as additive within the experimental error. As is seen in Table 3, 45.6% of the entries in Table 2 meet this criterion. This is still rather poor, as in the 65 intrascaffolding tests of Lu et al. (1997) 67.7% met the criterion. It is likely that intrascaffolding additivity is better than interscaffolding additivity.

Pro Variant Is Nonadditive. However, there is another explanation. In the CHYM column of Table 2, all 15 entries that do not involve P_1 Pro have a $\lvert\Delta G_I°\rvert$ of <400, i.e. that are additive within the experimental error. On the other hand, $\lvert\Delta G_I°\rvert$ for all six cases involving Pro is between 1500 and 1870 cal/mol, a substantial deviation from additivity. Inspection of the six rows involving Pro shows that they show strong nonadditivity, with the exception of some SGPA and SGPB entries. We therefore removed the Pro-involving entries and retabulated the results in Table 3. The fraction of remaining entries with a $\lvert\Delta G_I°\rvert$ of <400 cal/mol increased to 57.8%. More remarkably, $\lvert\Delta G_I°\rvert$ > 1200 cal/mol dropped from 17.5 to 4.2%. It is worth noting that the 65 intrascaffolding comparisons (Lu et al., 1997) listed in Table 3 do not involve Pro.

The removal of Pro may seem arbitrary. It is less arbitrary than it may appear. Pro is the only member of the coded set that is not an amino acid in a chemical sense. The Pro side chain differs from all the others by being attached to the main chain both at the CA atom (as all the others are) and at the N atom. Modeling shows that the Pro side chain is very difficult to fit into the S_1 cavities of all six enzymes that we study here. Additionally, all published three-dimensional structures of complexes of standard mechanism, canonical inhibitors with their cognate serine proteinases show the N(H) of the P_1 residue of the inhibitor donating a hydrogen bond to the enzyme. P_1 Pro does not have an available H atom and therefore cannot participate in such a bond. There are, as yet, no published three-dimensional structures of complexes where the P_1 residue is Pro.

However, K. Huang, W. Lu, S. Anderson, M. Laskowski, Jr., and M. N. G. James (manuscript in preparation), as a part of the X^{18}OMTKY3−SGPB complex study, for all X^{18} residues determined the high-resolution structure of the Pro^{18} OMTKY3−SGPB complex structure. In all other X^{18} OMTKY3−SGPB complexes examined thus far (Huang et al., 1995; Huang, 1995; Bateman et al., 1996), the nature of the X^{18} residue does not affect the conformation of the inhibitor's main chain and the nature of the interactions between the enzyme and the inhibitor's contact residues other than P_1. The Pro^{18} complex is an exception. In this complex, the inhibitor's main chain is slightly distorted and some of the interactions with SGPB at positions other than P_1 are altered. Such alterations might well differ in the two inhibitor scaffoldings and might entail different free energy costs, thus explaining nonadditivity. This argument is flawed by our finding that P_1 Pro is moderately additive for SGPB. Clearly, structures of complexes of P_1 Pro containing inhibitors in complex with other enzymes would be very helpful. We anticipate even greater main chain deviations in such complexes.

Further inspection of Table 2 shows that, even after Pro is removed, the HLE column contains especially many large $|\Delta G_I°|$ values. Bode et al. (1989) and Lu et al. (1997) remark on the much greater flexibility of the HLE S_1 cavity compared to that of the S_1 cavities of other serine proteinases. One of the arguments offered for this by Lu et al. (1997) is that HLE accounts for most of the observed large deviations from intrascaffolding additivity. The same conclusions apply here to interscaffolding additivity.

With Pro omitted, the interscaffolding additivity is still worse than the intrascaffolding additivity as might have been expected. However, both are quite good and clearly have a large predictive value in protein design.

Superimposition of Eglin c and OMTKY3 Complexes with Chymotrypsin. X-ray crystallography of many different X^{18} OMTKY3−SGPB complexes shows that for all X^{18}, except Pro^{18}, the nature of the P_1 residue affects only the P_1 side chain−S_1 cavity interaction; all other interactions are unaffected (Huang et al., 1995; Huang, 1995; K. Huang, K. S. Bateman, and M. N. G. James, unpublished; Bateman et al., 1996). Huang et al. (1995) and Lu et al. (1997) also showed that P_1 Gly can serve as an excellent standard and $\Delta\Delta G_a°$(Gly P_1 X) is a quantitative measure of the interaction of the side chain of the X residue with the S_1 cavity. From Table 1 for CHYM by analogy to eqs 2 and 3, we obtain $\Delta\Delta G_a°(\text{Gly } P_1 \text{ Leu})_{OMTKY3} = -6.15$ kcal/mol and $\Delta\Delta G_a°(\text{Gly } P_1 \text{ Leu})_{\text{eglin c}} = -6.18$ kcal/mol. $|\Delta G_I°| = 30$ cal/mol is an exceptionally small deviation. It seems clear that, in the two complexes, the P_1 Leu residue must bind to the S_1 cavity of chymotrypsin in a nearly identical manner. Fortunately, X-ray crystallographic data are available in the literature to make the comparison. Even though a large number of three-dimensional structures of enzyme−inhibitor complexes were determined (Bode & Huber, 1992), there are relatively few high-resolution structures of complexes of two inhibitors belonging to two different families but sharing the same P_1 residue and interacting with the same enzyme. The eglin c (Frigerio et al., 1992) and OMTKY3 (Fujinaga et al., 1987) complexes with CHYM are among the few.

The results are listed in Table 4. First, when all the CA atoms of the chymotrypsin molecules in the two complexes are superimposed, the resultant rmsd is 0.51 Å. This relatively large difference is presumably due to different crystal contacts as the two crystal forms are nonisomorphous. However, the environment in the enzyme−inhibitor contact region is much more similar. All 196 atoms in the 31 residues of the active site and S_1 pocket of chymotrypsin were superimposed except for the CE atom of Met^{192} as this methyl group shows an opposite orientation in both complexes. The rmsd is 0.24 Å. These 31 chymotrypsin residues are shown in a superimposition in Figure 5, where CHYM in the OMTKY3 complex is in yellow and CHYM in the eglin c complex is in red. The very small rmsd is not only a testimonial to the high quality of work in both X-ray laboratories but also a striking confirmation that the S_1 cavity in both complexes is virtually identical as expected from experimentally indistinguishable $\Delta\Delta G°$(Gly P_1 Leu) values. The similarity of the S_1 pockets can be even better appreciated in Figure 6, where the two pockets are shown separately. While the similarity is astonishingly good, two differences are readily seen. One of the water molecules in the OMTKY3 complex is not seen in the eglin c complex. As alluded to above, the CE atom of Met^{192} of chymotrypsin points toward P_1 Leu in the eglin c complex, while it points away from it in OMTKY3. Figure 6 also shows a part of both inhibitors. As the eight atoms of the P_1 Leu are of main interest, they were separately superimposed. As is seen in Table 4, the rmsd for these four atoms is 0.39 Å. This is quite small but clearly larger than the rmsd for the S_1 pockets.

Table 4: rmsds for the Superimposition of CHYM and Eglin c and CHYM and OMTKY3

atoms	rmsd (Å)
CHYM residues, C_α atoms	0.51 (238)
CHYM residues used for superimposition, all atoms except for CE of Met^{192}	0.24 (196)
P_1 Leu, all atoms	0.39 (8)

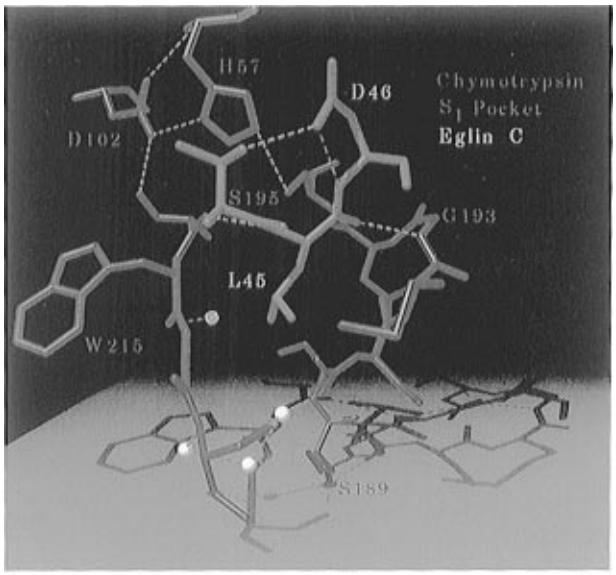
[a] Number of atoms indicated in parentheses.

On balance, the additivity of the P_1 Gly−P_1 Leu cycle in OMTKY3 and eglin c as determined by thermodynamic measurements is confirmed by X-ray crystallography.

Less Stringent Criteria. Additivity is the most stringent criterion for the similarity in behavior of relative equilibrium constants in two different systems. The finding in this paper that 45.6% of all P_1 combinations are strictly additive between eglin c and OMTKY3 is most impressive. Even more impressive is that 57.7% are strictly additive when Pro is omitted. On the other hand, 54.4 and 42.2%, respectively, are not. It is definitely of interest to use less stringent criteria.

Probably the least stringent is rank order. For chymotrypsin, this is L > I > S > G > E > P both for eglin c and for OMTKY3. Examination of Table 1 shows that it is not identical but nearly the same for the remaining five enzymes. Pro does not need to be omitted for this analysis. It is the worst in both the eglin c and OMTKY3 scaffoldings for CHYM, CARL, SGPA, and SGPB. It is poor but better than Asp and Glu for PPE and HLE in both scaffoldings.

Since one of the mutations we study is always a single change at P_1, the $-\Delta G_a°(P_1 = X)$ eglin c values can be plotted vs $-\Delta G_a°(P = X)_{OMTKY3}$. Figure 7 is a plot for chymotrypsin. Similar plots for five other enzymes are given



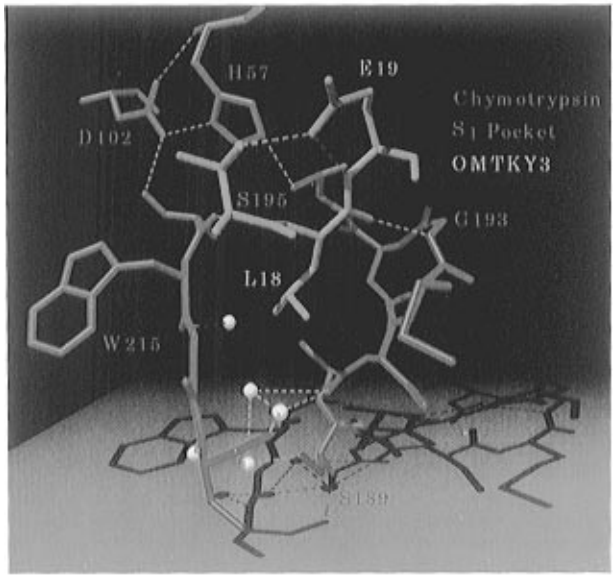


FIGURE 6: (a, top) Eglin c P_1 residue, Leu I45 (blue), bound to the S_1 pocket of chymotrypsin (red). (b, bottom) OMTKY3 P_1 residue, Leu I18 (green), bound to the S_1 pocket of chymptrypsin (red). Water molecules are shown as yellow spheres, and hydrogen bonds are drawn as dotted lines. This figure was made with RASTER3D (Bacon & Anderson, 1988).

in the Supporting Information. If the data in such plots were perfectly additive, all the points in such plots would lie precisely on a straight line of unit slope. For each plot, we provide the slope *s* which is a measure of sensitivity of the eglin c system compared to that of the OMTKY3 system to substitutions at P_1. We also provide the Pearson-product moment correlation coefficient, *r* (Campbell, 1989). $r = 1.00$ for data that are perfectly correlated, 0.00 for no correlation, and -1.00 for anticorrelation. In this case, *r* is a measure of how well the points fall on the lines. The *r* and *s* values for sets of all seven (or as many as could be measured) variants and for sets without Pro are listed in Table 5. It is seen there that *r* is very near the perfect unity for all cases, except perhaps PPE, and *s* rather is close to unity with several exceptions. This analysis restates all the previous conclusions. Changing P_1 in eglin c and in OMTKY3 produces closely similar relative effects.

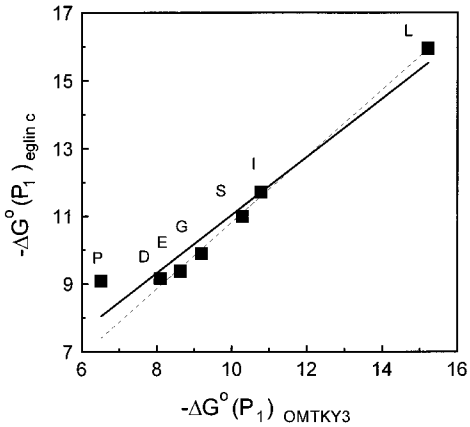


FIGURE 7: Correlation of the standard free energy change upon association of eglin c variants with chymotrypsin with the standard free energy change upon association of OMTKY3 variants with chymotrypsin. The solid line includes Pro, and the dashed line does not. Free energy changes are in kilocalories per mole.

Table 5: Correlation of the Standard Free Energy Change on Association of Eglin c Variants with Various Enzymes with the Standard Free Energy Change on Association of OMTKY3 Variants with the Same Enzymes[a]

	with Pro		without Pro	
	r	*s*	*r*	*s*
CHYM	0.97	0.86	1.00	0.98
PPE	0.94	0.63	0.91	1.03
CARL	0.99	0.89	0.99	0.98
SGPA	0.97	0.99	0.95	1.24
SGPB	0.98	1.17	0.98	1.55
HLE	0.95	0.94	0.98	1.03

[a] Figure 7 is an example.

DISCUSSION

Bigler et al. (1993) and Lu et al. (1997) measured K_a values for 25 different (5 noncoded, all 20 coded) P_1 variants of OMTKY3. They asked the following. Are their results general? A partial answer was obtained by comparing their $\Delta\Delta G^\circ$ values with those obtained in other avian ovomucoid third domains. One of the 12 additivity cycles they employed is shown in the top part of Figure 2. They showed that their system was strictly additive for 67.7% of the cycles examined and that $|\Delta G_I^\circ|$ was always less than 1.4 kcal/mol. This excellent additivity was, however, observed among inhibitors with remarkably similar amino acid sequences. They all belong to the same branch of the Kazal inhibitor family.

As already pointed out, standard mechanism, canonical inhibitors of serine proteinases consist of 20 different families. Each family is characterized by a common scaffolding that differs from that employed by other families. The scaffolding supports about a dozen residues that contact the enzyme in the proper orientation. Among these, the six or seven $[P_4-P_2'(P_3')]$ residues surrounding the reactive site share the common, canonical main chain conformation. Are the Bigler et al. (1993) and Lu et al. (1997) results applicable to the other inhibitor families?

Bigler et al. (1993) and Lu et al. (1997) answered the question by comparing their data to the literature data. The results are listed in the middle part of Table 4 of Lu et al. (1997) and compared in Table 5 of this paper. There was great difficulty in finding sufficiently comparable data. Different methods were used to measure K_a. Some of the

noncoded P_1 variants differed from ours. The measurements were often made at a temperature (37 °C) different than ours (21 °C) or most importantly at different pH values. Qasim et al. (1995) recently showed that for ionizable P1 residues (Asp, Glu, and Lys) the enzyme−inhibitor equilibrium constants show a large dependence on pH in the neutral and slightly alkaline pH region. Additionally, the enzyme sets with which the inhibitors were tested did not overlap well, reducing the number of useful comparisons. In view of all of the above, it is surprising that the *r* values are excellent and many *s* values are reasonably good.

In order to answer the question less ambiguously in this paper, we expressed seven P_1 variants of eglin c, an inhibitor from the potato I family, whose scaffolding is very different from that of OMTKY3. We measured the K_a's for these variants, whenever possible (see Results), with all six enzymes in the same laboratory, by the same measuring technique at the same pH, temperature, and ionic environment. We subjected the data to the most stringent comparison technique, additivity, and found them impressively additive especially when Pro is excluded. The exclusion of Pro seems reasonable, as it is the only member of the 25-residue set that is not an α-amino acid in a strict chemical sense. Therefore, while we studied only seven variants, we hope that interscaffolding additivity extends to 24 of the 25 residues studied by Lu et al. (1997), Pro being the only exception.

In combination with the data listed in Table 4 of Lu et al. (1997), our results make us optimistic about a claim that the results apply to all 20 protein families. However, this conclusion is not in accord with all of the published literature, especially if qualitative rather than quantitative data are considered. As an example, we cite the work of Kossiakoff et al. (1993), who found by phage display that among the coded P_1 variants of bovine pancreatic trypsin inhibitor (Kunitz), BPTI, P_1 Asn and His are best for the inhibition of chymotrypsin. In sharp contrast, Lu et al. (1997) report that in OMTKY3 P_1 Tyr, Trp, and Phe are best, followed by Leu and Met. If the results of both groups are correct, OMTKY3 and BPTI are not interscaffolding additive for the inhibition of chymotrypsin.

We proposed generalizing (with some caveats) the results to all L-α-amino acid and glycyl residues and to all standard mechanism, canonical protein inhibitor families. Another generalization is of great interest. Are substitutions at positions other than P_1 but in the canonical region of the combining loop also interscaffolding additive? The available data seem too scarce to deal with this question now.

Concerning P_1 Pro, for every one of the six enzymes tested, P_1 Pro is locally deleterious in both eglin c and OMTKY3 scaffoldings. However, for every enzyme, it is much less deleterious in eglin c than in OMTKY3. It seems likely that the presence of a disulfide bridge at P_3 of OMTKY3 makes its main much harder to deform than the eglin c main chain, thus causing an additional problem in binding of P_1 Pro.

Eglin c as an Inhibitor. As already pointed out, we did not expect and did not find K_a's for the P_1 = Leu wild types of eglin c and of OMTKY3 to be the same. Eglin c inhibits five of the six enzymes we study more strongly than OMTKY3. It inhibits PPE much more weakly. These results raise to questions. One is about the molecular explanation of the differences. Unfortunately, the only cases for which there are three-dimensional structures of both eglin c and OMTKY3 interacting with the same enzyme are their complexes with chymotrypsin. While eglin c inhibits more strongly, the difference is so slight that it is difficult to attempt an explanation. Dramatic differences are seen for SGPA and SGPB on one hand and for PPE on the other.

The second question is a practical one. In suggesting potential inhibitors to try for newly characterized enzymes, it is best to use eglin c mutants if the enzyme shows a substrate specificity similar to that of SGPA and SGPB and OMTKY3 mutants if the enzyme shows a substrate specificity similar to that of PPE. In a few recent studies in our laboratory, we have adopted such a strategy. It was successful, but until a molecular explanation is provided, this is largely anecdotal.

ACKNOWLEDGMENT

We thank Mark Bauer and Yiping Sun for mass spectrometry analysis, Dan Toerner and Ellen Wang for DNA sequence analysis, Jim Armpriester for help with eglin c purification, and Christy Olson for DNA manipulations. D. A. Estell, the late M. Laskowski, Sr., L. Smillie, and J. Travis contributed gifts of very pure enzyme samples which were used as reference standards. M. Tashiro, S. J. Park, M. Wieczorek, R. Wynn, and W. Lu contributed to our thinking about intrascaffolding and interscaffolding additivity.

SUPPORTING INFORMATION AVAILABLE

Five figures of $-\Delta G°(P_1)_{\text{eglin c}}$ vs $-\Delta G°(P_1)_{\text{OMTKY3}}$, analogous to Figure 7, for PPE, CARL, SGPA, SGPB, and HLE (6 pages). Ordering information is given on any current masthead page.

REFERENCES

Ackers, G. K., & Smith, F. R. (1985) *Annu. Rev. Biochem. 54*, 597−629.

Ardelt, W., & Laskowski, M., Jr. (1985) *Biochemistry 24*, 5313−5320.

Bacon, D. J., & Anderson, W. F. (1988) Raster3D. A fast algorithm for rendering space-filling molecule pictures, *J. Mol. Graphics 6*, 219−220.

Bateman, K. S., Huang, K., Lu, W., Anderson, S., Laskowski, M., Jr., & James, M. N. G. (1996) Poster presented at Keystone Symposium on Proteolytic Enzymes and Inhibitors in Biology and Medicine, March 25−31, 1996, Keystone, CO.

Bernstein, F. C., Koetzle, T. F., Williams, G. J., Meyer, E. E., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., & Tasumi, M. (1977) PDB. The Protein Data Bank: a computer-based archival file for macromolecular structures, *J. Mol. Biol. 112*, 535−542.

Bidlingmeyer, U. D. V., Leary, T. R., & Laskowski, M., Jr. (1972) *Biochemistry 11*, 3303−3310.

Bigler, T., Lu, W., Park, S. J., Tashiro, M., Wieczorek, M., Wynn, R., & Laskowski, M., Jr. (1993) *Protein Sci. 2*, 786−799.

Biosym Technologies (1993) *InsightII, version 2.3.0*, San Diego.

Bode, W., & Huber, R. (1992) *Eur. J. Biochem. 204*, 433−451.

Bode, W., Papamokos, E., & Musil, D. (1987) *Eur. J. Biochem. 166*, 673−692.

Bode, W., Meyer, E., & Powers, J. C. (1989) *Biochemistry 28*, 1951−1963.

Campbell, R. C. (1989) *Statistics for Biologists*, 3rd ed., pp 316−320, Cambridge University Press, Cambridge.

Carter, P. J., Winter, J., Wilkinson, J. J., & Fersht, A. R. (1984) *Cell 38*, 835−840.

Collaborative Computational Project, Number 4 (1994) The CCP4 Suite: Programs for Protein Crystallograpy, *Acta Crystallogr. D50*, 760−763.

Dauter, Z., & Betzel, C. (1991) *Acta Crystallogr. B47*, 707−730.

Empie, M. W., & Laskowski, M., Jr. (1982) *Biochemistry 21*, 2274−2284.
Finkenstadt, W. R., Hamid, M. A., Mattis, J. A., Schrode, J., Sealock, R. W., Wang, D., & Laskowski, M., Jr. (1974) *Bayer-Symp. 5*, 389−411.
Frigerio, F., Coda, A., Pugliese, L., Lionetti, C., Menegatti, E., Amiconi, G., Schnebli, H. P., Ascenzi, P., & Bolognesi, M. (1992) *J. Mol. Biol. 225*, 107−123.
Fujinaga, M., Sielecki, A. R., Read, R. J., Ardelt, W., Laskowski, M., Jr., & James, M. N. G. (1987) *J. Mol. Biol. 195*, 397−418.
Green N. M., & Work, E. (1953) *Biochem. J. 54*, 347−352.
Gros, P., Betzel, Ch., Dauter, Z., Wilson, K. S., & Hol, W. G. J. (1989) *J. Mol. Biol. 210*, 347−367.
Heinz, D. W., Priestle, J. P., Rahuel, J., Wilson, K. S., & Grütter (1991) *J. Mol. Biol. 217*, 353−371.
Huang, K. (1995) Structural studies of the interactions between serine proteinases and protein inhibitors, Ph.D. Thesis, University of Alberta, Edmonton, Alberta, Canada.
Huang, K., Lu, W., Anderson, S., Laskowski, M., Jr., & James, M. N. G. (1995) *Protein Sci. 4*, 1985−1997.
Kato, I., Schrode, J., Kohr, J., & Laskowski, M., Jr. (1987) *Biochemistry 26*, 193−201.
Kossiakoff, A. A., Hynes, T., & de Vos, A. (1993) *Biochem. Soc. Trans. 21*, 614−618.
Laskowski, M., Jr. (1986) in *Nutritional and toxicological significance of enzyme inhibitors in food* (Friedman, M., Ed.) pp 1−17, Plenum Press, New York.
Laskowski, M., Jr., & Sealock, R. W. (1971) *Enzymes 3*, 375−473.
Laskowski, M., Jr. & Kato, I. (1980) *Annu. Rev. Biochem. 49*, 593−626.
Laskowski, M., Jr., Tashiro, M., Empie, M. W., Park, S. J., Kato, I., Ardelt, W., & Wieczorek, M. (1983) in *Proteinase Inhibitors: Medical and Biological Aspects* (Katunama, N., et al., Eds.) pp 55−68, Japan Scientific Society Press and Springer-Verlag, Tokyo and Berlin.
Laskowski, M., Jr., Kato, I., Ardelt, W., Cook, J., Denton, A., Empie, M. W., Kohr, W. J., Park, S. J., Parks, K., Schatzley, B. L., Schoenberger, O. L., Tashiro, M., Vichot, G., Whatley, H. E., Wieczorek, A., & Wieczorek, M. (1987) *Biochemistry 26*, 202−221.
Leary, R., & Laskowski, M., Jr. (1973) *Fed. Proc. 32*, 456.
Lu, W., Apostol, I., Qasim, M. A., Warne, N., Wynn, R., Zhang, W. L., Anderson, S., Chiang, Y. W., Ogin, E., Rothberg, I., Ryan, K., & Laskowski, M., Jr. (1997) *J. Mol. Biol.* (in press).
McPhalen, C. A., & James, M. N. G. (1988) *Biochemistry 27*, 6582−6598.
Nicholls, A., Sharp, K., & Honig, B. (1991) Grasp. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons, *Proteins: Struct., Funct., Genet. 11*, 281−296.
Park, S. J. (1985) Effect of amino acid replacements in ovomucoid third domains upon their association with serine proteinases, Ph.D. Thesis, Purdue University, West Lafayette, IN.
Qasim, M. A., Ranjbar, M. R., Wynn, R., Anderson, S., & Laskowski, M., Jr. (1995) *J. Biol. Chem. 270*, 27419−27422.
Read, R. J., & James, M. N. G. (1986) In *Proteinase Inhibitors* (Barrett, A. J., & Salvesen, G., Eds.) pp 301−336, Elsevier Science Publishers B. V., Amsterdam.
Rink H., Liersch, M., Sieber, P., & Meyer, F. (1984) *Nucleic Acids Res. 12*, 6369−6387.
Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989) *Molecular Cloning, A Laboratory Manual*, Cold Springer Harbor Laboratory Press, Plainview, NY.
Saunders, W. W., Pedroni, J. A., & Monahan, P. M. (1991) *Gene 102*, 277−282.
Schecter, I., & Berger, A. (1967) *Biochem. Biophys. Res. Commun. 27*, 157−162.
Sealock, R. W., & Laskowski, M., Jr. (1969) *Biochemistry 8*, 3703−3710.
Seemüller, U., Eulitz, M., Fritz, H., & Strobl, A. (1980) *Z. Physiol. Chem. 361*, 1841−1846.
Vasantha, N., Thompson, L. D., Rhodes, C., Banner, C., Nagle, J., & Filpula, D. (1984) *J. Bacteriol. 159*, 811−819.
Vincent, J.-P., & Lazdunski, M. (1972) *Biochemistry 11*, 2967−2977.
Wells, J. A. (1990) *Biochemistry 29*, 8509−8517.
Wynn, R. (1990) Design of a specific human leukocyte elastase inhibitor based on ovomucoid third domains, Ph.D. Thesis, Purdue University, West Lafayette, IN.

BI9620870