due to the controls scoring lower than normal during block 5 but scoring higher than normal for block 6, whereas the reverse was true for the SIV-infected monkeys for the same blocks. In any event, the interaction cannot account for the pattern of cognitive and motor impairments. The tolerance limit method yielded no consistent findings.

16. J. M. Ward *et al.*, *Am. J. Pathol.* **127**, 199 (1987).
17. L. R. Sharer *et al.*, *J. Med. Primatol.* **20**, 211 (1991).

18. With few exceptions, the monkeys displayed food intake, body temperatures, and body weights that were in the normal range; M007, M011, and M031 were occasionally febrile, and M007 and M014 showed occasional, brief (1 to 2 days) losses of appetite during the course of training after inoculation, but these episodes generally postdated, and so cannot account for, the cognitive and motor deficits. For the most part, the SIV-infected monkeys succumbed to diarrhea, wasting, or opportunistic infection only very late in the course of disease, about a week before they became moribund and were euthanized (Table 2).

19. M. Perdices and D. A. Cooper, *Ann. Neurol.* **25**, 460 (1989); E. N. Miller *et al.*, *Neurology* **40**, 197 (1990); F. L. Wilkie *et al.*, *Arch. Neurol.* **47**, 433 (1990).
20. R. W. Price *et al.*, *Science* **239**, 586 (1988); C. A. Wiley *et al.*, *Ann. Neurol.* **29**, 651 (1991).

21. We thank P. Brouwers, P. Pizzo, M. Mishkin, and M. Brownstein for their support in all phases of this study; M. Murphey-Corb for both the Delta B670 strain of SIV_{smm} and the rabbit polyclonal antiserum to SIV; L. Martin for numerous discussions; J. Parkinson and K. Pettigrew for assistance with the statistical analyses; R. Byrum and T. Moskal for help conducting the experiments; and R. Byrum additionally for invaluable assistance coordinating and supervising the various aspects of the study.

4 October 1991; accepted 23 December 1991

Converting Trypsin to Chymotrypsin: The Role of Surface Loops

LIZBETH HEDSTROM, LASZLO SZILAGYI, WILLIAM J. RUTTER*

Trypsin (Tr) and chymotrypsin (Ch) have similar tertiary structures, yet Tr cleaves peptides at arginine and lysine residues and Ch prefers large hydrophobic residues. Although replacement of the S1 binding site of Tr with the analogous residues of Ch is sufficient to transfer Ch specificity for ester hydrolysis, specificity for amide hydrolysis is not transferred. Trypsin is converted to a Ch-like protease when the binding pocket alterations are further modified by exchange of the Ch surface loops 185 through 188 and 221 through 225 for the analogous Tr loops. These loops are not structural components of either the S1 binding site or the extended substrate binding sites. This mutant enzyme is equivalent to Ch in its catalytic rate, but its substrate binding is impaired. Like Ch, this mutant utilizes extended substrate binding to accelerate catalysis, and substrate discrimination occurs during the acylation step rather than in substrate binding.

CHYMOTRYPSIN (CH) AND TRYPSIN (Tr) have extensive sequence identity (*1*) and seemingly superimposable main chain structures (*2, 3*) yet have very different substrate specificities. Until recently, the substrate specificity of these pancreatic serine proteases was believed to be a simple function of the steric and electrostatic properties of the S1 binding site (*2, 4*). The S1 binding sites of Tr and Ch are nearly identical in structure and primary sequence (Figs. 1 and 2). Chymotrypsin has a hydrophobic S1 binding pocket formed by residues 189 through 195, 214 through 220, and 225 through 228 (*5*); this feature ostensibly explains Ch's specificity for large hydrophobic residues. The preference of Tr for Lys and Arg results from the presence of Asp^{189} (Ser in Ch) at the bottom of the S1 binding pocket (*6, 7*). The structural basis of substrate specificity in these enzymes is more complex than these simple mechanistic postulates would imply. Mutation of Asp^{189} to Ser (D189S) in Tr does not switch the substrate specificity of Tr to that of Ch but creates a poor, nonspecific protease (*8, 9*) (Table 1). We have further elucidated the structural determinants of specificity and activity in Tr and Ch by showing that the S1 binding pocket determines the specificity of ester hydrolysis, whereas specific amide hydrolysis requires both the proper S1 binding site and more distal binding site interactions. These interactions are profoundly influenced by surface loops that do not directly contact the substrate.

Ester hydrolysis is intrinsically less specific than amide hydrolysis. The hydrolysis of hydrophobic oligopeptide amide substrates by Tr is 10^5-fold less efficient than by Ch, but only a 10- to 100-fold difference in k_{cat}/K_m (k_{cat}, catalytic rate constant; K_m, Michaelis constant) exists between Tr and Ch for the hydrolysis of *N*-acetylphenylalaninyl-*p*-nitrophenylester (AcF-pNP) and succinyl-AlaAlaProPhe-thiobenzyl ester (sucAAPF-SBzl) (Tables 1 and 2). The k_{cat} values are similar, whereas K_m differs 10- to 100-fold. This apparent lack of specificity for ester hydrolysis may be a consequence of k_{cat}/K_m approaching the rate of diffusion, which would effectively limit selectivity (*10*). The steady-state kinetic constants for the hydrolysis of both AcF-pNP and sucAAPF-SBzl by D189S are equivalent to those of Ch, indicating that the Asp^{189} to Ser mutation in Tr is sufficient to change the specificity of ester hydrolysis.

Unlike ester hydrolysis, specific amide hydrolysis by pancreatic serine proteases is influenced by the length of the oligopeptide substrate, and specificity is largely determined by the extended substrate binding sites (*11*). The steady-state kinetic parameters for the hydrolysis of single amino acid and oligopeptide amide substrates by Ch, Tr, and D189S are compared in order to assess the contribution to amide substrate specificity of the S1 binding site and the extended binding sites (Table 2). Chymotrypsin hydrolyzes the oligopeptide amide substrates succinyl-AlaAlaProPhe-7-amino-4-methylcoumarin (sucAAPF-AMC) and succinyl-AlaAlaProPhe-*p*-nitroanilide (sucAAPF-pNA) 10^5-fold faster than acetylphenylalaninamide (AcF-NH_2), as measured from k_{cat}/K_m. Trypsin hydrolyzes sucAAPF-AMC only tenfold faster than AcF-NH_2; thus, Tr hydrolyzes sucAAPF-AMC 10^5-fold more slowly than Ch. This difference in catalytic activity for extended oligopeptide substrates contrasts dramatically with the modest differences between Tr and Ch in AcF-NH_2 hydrolysis. Trypsin cannot utilize extended substrate binding to accelerate hydrolysis of Ch-specific substrates. Clearly, the extended binding sites contribute substantially to substrate specificity. Efficient hydrolysis of amides requires both the correct P1 residue and an extended substrate (*11*). The specificity-determining transition state must include the substrate P2, P3, and P4 residues, as well as the P1 residue.

The hydrolysis of AcF-NH_2 was also used to probe the function of the S1 binding pocket independent of the extended binding sites. Trypsin hydrolyzes AcF-NH_2 100-fold more slowly than Ch (k_{cat}/K_m, Table 2). The k_{cat} value is 10^3-fold less than that of Ch; surprisingly, K_m is 10-fold lower. Perhaps Tr binds AcF-NH_2 nonproductively. The D189S mutant Tr partially reconstructs Ch amide specificity. The k_{cat} for hydrolysis of AcF-NH_2 is improved 50- to 100-fold to 5% of the Ch value. However, K_m is also increased 100-fold over that of Tr, so k_{cat}/K_m does not change. Thus, the S1 binding site of D189S is significantly compromised relative to Ch.

The complete replacement of the S1 bind-

L. Hedstrom and W. J. Rutter, Hormone Research Institute and Department of Biochemistry and Biophysics, University of California, San Francisco, CA 94143-0534.
L. Szilagyi, Biochemistry Department, Eötvös Loránd University, Budapest, Hungary.

*To whom correspondence should be addressed.

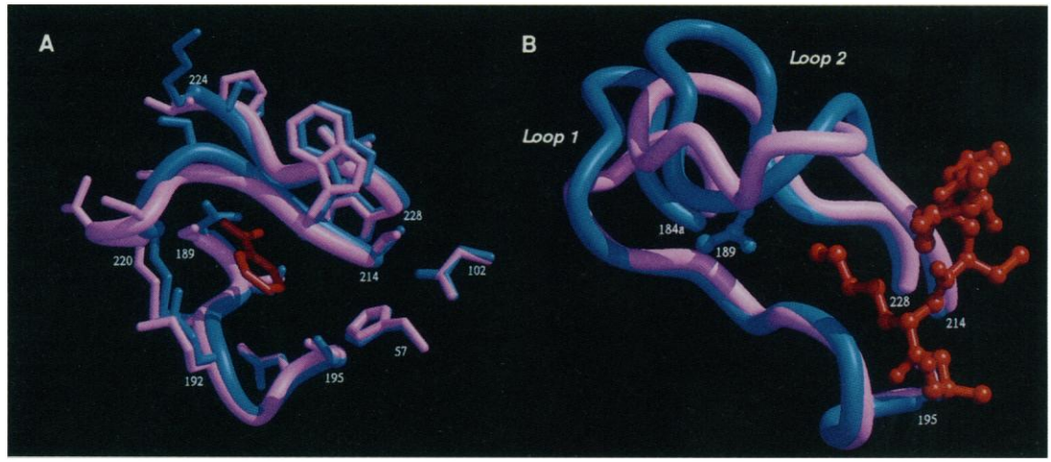

Fig. 1. (**A**) The Tr and Ch binding pockets. The structures of the bovine Tr-benzamidine complex (Brookhaven Protein Data Bank identification number 3ptb) (*30*) and bovine γ-Ch (2gch) (*5*) have been aligned at the catalytic triad residues His^{57}, Asp^{102}, and Ser^{195} and at the oxyanion hole (Gly^{193} and Ser^{195}). Trypsin is blue, Ch is pink, and benzamidine is red. The binding pocket residues 189 through 195, 214 through 220, and 224 through 228 are shown, as well as the catalytic triad. (**B**) The location of loops 1 and 2 in relation to the substrate binding sites. The Tr-bovine pancreatic trypsin inhibitor complex (BPTI) (2ptc) (*31*) and bovine γ-Ch are aligned as in (A). Trypsin is blue, Ch is pink, and BPTI residues 12I through 16I are red. BPTI residue Lys^{15I} is in the S1 binding pocket, and residues 12I through 14I form an antiparallel β sheet with Tr residues 214 through 218. The side chains of Tr Asp^{189} and Ser^{195} are also shown. This figure was produced with the MidasPlus software system from the Computer Graphics Laboratory, University of California, San Francisco, and with software written by T. Hynes (*32*).

ing pocket was not sufficient to transfer amide substrate specificity. The substrate P1 residue contacts residues 189 through 192, 214 through 216, and 225 through 227 (*5*), of which only three residues differ. The Ser^{189} and Met^{192} residues in Ch are Asp^{189} and Gln^{192} in Tr while residue 219 is deleted in Tr. In addition, Ser^{190} in Ch is rotated out of the binding pocket by hydrogen bonding with Thr^{138}. In Tr, residue 138 is Ile, and Ser^{190} extends into the binding pocket, where it interacts with substrates (*6*). Trypsin mutants with Ch-like S1 binding pockets were constructed with site-directed mutagenesis (*12*): Tr→Ch[S1−1] contains D189S, Q192M, and I138T mutations; Tr→Ch[S1−2] contains D189S with Thr inserted at position 219; and Tr→Ch[S1−3] contains all four alterations: D189S, Q192M, I138T, and 219T (Fig. 2). Table 1 presents the k_{cat}/K_m values for the reaction of Ch, Tr, D189S, and the binding-pocket mutants with a series of fluorogenic oligopeptides with varying P1 residues. All of the Tr mutants bind quantitatively to soybean trypsin inhibitor (SBTI) Sepharose. Thus, the enzymes appear to be properly folded and active. None of the binding pocket mutants are significantly better proteases than D189S. Therefore, determinants of catalytic specificity must exist outside the S1 binding pocket.

The residues that connect the walls of the binding pocket, 185 through 188 (loop 1, residues 184a through 188a in Tr) and 221 through 225 (loop 2), are quite different (Figs. 1 and 2). Loop 1 is highly conserved in Ch; a similar but distinct loop is equally conserved in Tr (Fig. 2). Loop 2 in Ch is highly conserved and distinct from the Tr sequences, which are variable (*13*). These surface loops do not contact the substrate in any of these enzymes. Despite this fact, the conservation of these sequences in the respective proteases suggests that they may have a role in determining substrate specificity.

Trypsin mutants that contain the Ch-like S1 binding pocket (Tr→Ch[S1−3]) and Ch-like loop 1 (Tr→Ch[S1+L1]) or Ch loop 2 (Tr→Ch[S1+L2]) or both loops (Tr→Ch[S1+L1+L2]) were constructed by site-directed mutagenesis (*12*). All Tr mutants bound quantitatively to SBTI Sepharose, suggesting that the enzymes are properly folded and active. Tr→Ch[S1+L1] displayed poor Ch-like activity, although Tr→Ch[S1+L2] was modestly (fivefold) more active than D189S (Table 1). The combination of both loops in Tr→Ch[S1+L1+L2] increased Ch-like activity several 100-fold. The k_{cat}/K_m values for the hydrolysis of the Phe, Tyr, Leu, and Trp substrates by Tr→Ch[S1+L1+L2] are within 10^2- to 10^3-fold of Ch (Table 1). The k_{cat} values for the hydrolysis of sucAAPF-AMC and sucAAPF-pNA by Tr→Ch[S1+L1+L2] are comparable to those of Ch (Table 2). Tr→Ch[S1+L1+L2] hydrolyzed hydrophobic substrates 10^2- to 10^3-fold faster than the Lys substrate and discriminates between Lys and hydrophobic substrates equivalently to Ch (Table 1). Therefore, loops 1 and 2 interact synergistically to determine substrate specificity. The magnitude of k_{cat}/K_m for Tr→Ch[S1+L1+L2]–catalyzed hydrolysis is comparable to that of mammalian Ch-like proteases such as cathepsin G and chymase (*14*). Like D189S, Tr→Ch[S1+L1+L2] hydrolyzes esters as well as Ch does (Table 2).

The steady-state kinetic constants for the hydrolysis of AcF-NH_2 are similar for Tr→Ch[S1+L1+L2] and D189S. Although the amino acid sequence of the S1 binding pocket of Tr→Ch[S1+L1+L2] is identical to that in Ch, the function of the S1 binding pocket is still imperfect. Tr→Ch[S1+L1+L2], like Ch, hydrolyzes oligopeptide amides 10^4-fold better than AcF-NH_2. Thus, Tr→Ch[S1+L1+L2] re-establishes the contribution of the extended

```
                                  138  182       190       200       210       220       230
                                   |    |         |         |         |         |         |
Cow chymotrypsin A* (Ch)           t    cag-asgv-sscmgdsggplvckkngawtlvgivswgsstcs-tstpgvyarv
Dog chymotrypsin†                  t    cag-asgv-sscmgdsggplvcqkdgvwtlagivswgsgtcs-tstpgvysrv
Cow trypsin‡                       i    cagyleggkdscqgdsggpvvc--sgk--lqgivswgsg-caqknkpgvytkv
Rat trypsin-II§ (Tr)               i    cvgfleggkdscqgdsggpvvc--nge--lqgivswgyg-calpdnpgvytkv
D189S                              i    cvgfleggkSscqgdsggpvvc--nge--lqgivswgyg-calpdnpgvytkv
Tr→Ch[S1-1]                        T    cvgfleggkSscMgdsggpvvc--nge--lqgivswgyg-calpdnpgvytkv
Tr→Ch[S1-2]                        i    cvgfleggkSscqgdsggpvvc--nge--lqgivswgygTcalpdnpgvytkv
Tr→Ch[S1-3]                        T    cvgfleggkSscMgdsggpvvc--nge--lqgivswgygTcalpdnpgvytkv
Tr→Ch[S1+L1]                       T    cvg-ASgG-SscMgdsggpvvc--nge--lqgivswgygTcalpdnpgvytkv
Tr→Ch[S1+L2]                       T    cvgfleggkSscMgdsggpvvc--nge--lqgivswgSgTcS-TSTpgvytkv
Tr→Ch[S1+L1+L2]                    T    cvg-ASgG-SscMgdsggpvvc--nge--lqgivswgSgTcS-TSTpgvytkv
                                        ------                                      -----
                                        Loop 1                                      Loop 2
```

Fig. 2. Trypsin mutants. The site-directed mutagenesis method of Kunkel was used to create Tr mutants (*12*). Mutant enzymes were expressed in *Escherichia coli* with the pTRAP vector (*26*) and in the yeast expression system described for carboxypeptidase A1 (*25*). Chymotrypsin numbering is used. Capital letters denote mutations, and bold letters identify the amino acid residues that form the S1 binding pocket. Loops 1 and 2 are denoted with dashed lines. *(*33*). †(*34*). ‡(*35*). §(*36*).

Table 1. Hydrolysis of sucAAP-X-AMC by Tr mutants. The hydrolysis of sucAAP-X-AMC, where X is Phe, Tyr, Trp, Leu, and Lys, was monitored spectrofluorimetrically [excitation at 380 nm and emission at 460 nm (*27*)] in 0.1 M NaCl, 10 mM $CaCl_2$, 50 mM Hepes buffer, at pH 8.0 and 37°C, at substrate concentrations $\ll K_m$. Under these conditions k_{cat}/K_m is determined from the slope of the plot of rate versus substrate concentration. Data were analyzed with standard linear regression analysis and the k_{cat}/K_m values are the average of at least two experiments (SEM < 15% in all cases). Specificity is determined by the ratio of k_{cat}/K_m(hydrophobic substrates) to k_{cat}/K_m(Lys substrate). Substrates were prepared as described (*8*). Bovine Ch (*N*α-tosyl-L-lysine-chloromethylketone–inactivated) was purchased from Sigma Chemical Co.

Enzyme	k_{cat}/K_m (M^{-1} s^{-1})					Specificity
	Phe	Tyr	Trp	Leu	Lys	
Ch	1.6×10^6	4.5×10^6	6.8×10^6	1.2×10^5	850	10^2 to 10^3
Tr	4.5*	1.8*	0.2*	0.2*	1.2×10^6†	10^{-6} to 10^{-7}
D189S	33*	150*	2.3*	4.7*	16†	0.1 to 10
Tr→Ch[S1−1]	80	83	2.8			
Tr→Ch[S1−2]	47					
Tr→Ch[S1−3]	67	720	58	17		
Tr→Ch[S1+L1]	17	67	3.3	3.3		
Tr→Ch[S1+L2]	150	750		113		
Tr→Ch[S1+L1+L2]	2.8×10^3	2.0×10^4	2.0×10^3	1.0×10^3	34	10^2 to 10^3

*(*28*). †(*8*).

Table 2. Steady-state kinetic constants for Ch, Tr, and Tr mutants for the hydrolysis of AcF-pNP, sucAAPF-SBzl, AcF-NH_2, sucAAPF-AMC, and sucAAPF-pNA. Conditions are as described in Table 1. The sucAAPF-pNA and sucAAPF-SBzl hydrolysis was monitored spectrophotometrically (*18*). The AcF-pNP hydrolysis was measured spectrophotometrically (extinction coefficient at 400 nm $\varepsilon_{400} = 5.4$ mM^{-1} cm^{-1} at pH 6.5). The AcF-NH_2 hydrolysis was measured as described (*29*). Data were analyzed with KinetAsyst software and values are the average of at least two experiments. The Ch values are in good agreement with other reports (*15, 18, 29*). The compounds sucAAPF-pNA, sucAAPF-SBzl, AcF-pNP, and AcF-NH_2 were purchased from Sigma Chemical Co.

Enzyme	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m (M^{-1} s^{-1})
	AcF-pNP, pH 6.5†*		
Ch	30 ± 3	50 ± 15	(6 ± 1) × 10^5
Tr	29.8 ± 0.7	650 ± 80	(4.6 ± 0.1) × 10^4
D189S	40 ± 10	70 ± 20	(6.0 ± 0.6) × 10^5
Tr→Ch[S1+L1+L2]	33 ± 3	27 ± 8	(1.2 ± 0.1) × 10^6
	sucAAPF-SBzl		
Ch	52 ± 2	16 ± 1	(3.1 ± 0.3) × 10^6
Tr	36 ± 1	470 ± 35	(7.0 ± 0.2) × 10^4
D189S	33 ± 5	27 ± 3	(1.23 ± 0.03) × 10^6
Tr→Ch[S1+L1+L2]	36 ± 5	21 ± 3	(1.7 ± 0.1) × 10^6
	*sucAAPF-SBzl, pH 6.5**		
Ch	29 ± 1	12 ± 1	(2.5 ± 0.2) × 10^6
Tr	9.4 ± 0.7	115 ± 15	(8.5 ± 0.5) × 10^4
D189S	23 ± 1	85 ± 15	(3.0 ± 0.7) × 10^5
Tr→Ch[S1+L1+L2]	47 ± 3	80 ± 10	(5.7 ± 0.6) × 10^5
	AcF-NH_2		
Ch	0.43 ± 0.08	(2.3 ± 0.08) × 10^4	18 ± 3
Tr‡	(2.2 ± 0.6) × 10^{-4}	1000 ± 200	0.2 ± 0.1
D189S	0.018 ± 0.006	(1.7 ± 0.1) × 10^5	0.10 ± 0.01
Tr→Ch[S1+L1+L2]	0.018 ± 0.003	(6.5 ± 0.2) × 10^4	0.32 ± 0.08
	sucAAPF-AMC		
Ch	38 ± 2	12 ± 3	(3.3 ± 0.3) × 10^6
Tr§	(6.3 ± 0.3) × 10^{-3}	780 ± 60	8
D189S§	0.042 ± 0.001	690 ± 30	62
Tr→Ch[S1+L1+L2]‖	≥25	≥5000	(2.8 ± 0.2) × 10^{-3}
	sucAAPF-pNA		
Ch	49 ± 2	86 ± 5	(5.6 ± 0.5) × 10^5
Tr			
D189S	0.29 ± 0.01	15,000 ± 2,000	21 ± 1
Tr→Ch[S1+L1+L2]	13.0 ± 0.5	7,300 ± 300	(1.7 ± 0.1) × 10^3

*Assays were done in 0.1 M NaCl and 50 mM NaH_2PO_4, at pH 6.5 and 37°C. †AcF-pNP is a D,L mixture; kinetic parameters are corrected for content of L-AcF-pNP, with the assumption that the D isomer does not interfere. ‡No hydrolysis of AcF-NH_2 was observed with a mutant Tr in which the active site Ser^{195} is replaced with Ala, which suggests that the activity is specific to wild-type Tr and not due to a contaminating protease. §From (*8*). ‖No saturation is observed at 1.0 mM sucAAPF-AMC.

binding sites to substrate specificity. The k_{cat} values for hydrolysis of sucAAPF-AMC and sucAAPF-pNA by Tr→Ch[S1+L1+L2] are comparable to those of the Ch-catalyzed reactions; however, the K_m values are much higher. Because Tr→Ch[S1+L1+L2] has a catalytic turnover similar to that of Ch, it must also form an extended transition state structure similar to that of Ch.

Serine proteases hydrolyze amides and esters by a three-step mechanism of substrate binding (K_s), acylation of Ser^{195} (k_2), and hydrolysis of the acylenzyme intermediate (deacylation, k_3) (see Table 3 for scheme). The mechanistic kinetic constants for the hydrolysis of sucAAPF-AMC, sucAAPF-pNA, and AcF-NH_2 by Ch, Tr, D189S, and Tr→Ch[S1+L1+L2] can be calculated from the data in Table 2 and are shown in Table 3 (*15*). Classic experiments from many laboratories have established that acylation is rate determining for the hydrolysis of small amide substrates [reviewed in (*16*)]. However, it is not widely recognized that deacylation can be rate determining for oligopeptide amide substrates (*17*). The similarity of k_{cat} values for the hydrolysis of sucAAPF-pNA and sucAAPF-SBzl by Ch suggests that deacylation is also rate determining for the hydrolysis of specific oligopeptide amides by Ch (*18*). The k_{cat} values for the Ch-catalyzed hydrolysis of sucAAPF-SBzl, sucAAPF-AMC, and sucAAPF-pNA are equivalent (Table 2), indicating that the rate determining step for these substrates does not involve the leaving groups. Therefore, deacylation must be rate determining for oligopeptide amide as well as ester substrates. This result has been substantiated with a nucleophile trapping experiment (*19*). The k_{cat} values (Table 2) for the hydrolysis of sucAAPF-SBzl, sucAAPF-AMC, and sucAAPF-pNA by Tr→Ch[S1+L1+L2] are similar. Thus, as in Ch, the deacylation step is rate determining for Tr→Ch[S1+L1+L2]–catalyzed hydrolysis of oligopeptides.

Because k_3 is similar and independent of substrate length for Ch, Tr, and the two mutant enzymes, specificity must be largely determined in the binding and acylation steps. This arrangement is biologically preferable: an enzyme that undergoes covalent modification should make as few mistakes as possible (that is, discriminate in binding and acylation) but should be able to process the occasional mistake to avoid inactivation (and thus little selectivity in deacylation). Acylation appears to be the more discriminating step. Only a 10- to 100-fold difference in the binding of amide substrates exists between Tr and Ch (Table 3). In fact, Tr actually binds AcF-NH_2 better than Ch. The *p*-nitrophenyl and thiobenzyl esters are only 10- to 100-fold better substrates for Ch

Table 3. Mechanistic kinetic parameters for Ch, Tr, and Tr mutants for amide hydrolysis. The mechanistic kinetic parameters for the hydrolysis of AcF-NH_2, sucAAPF-AMC, and sucAAPF-pNA were derived from the steady-state kinetic constants of Table 2 according to the mechanism

$$E + RCOX \underset{}{\overset{K_s}{\rightleftharpoons}} E{\cdot}RCOX \xrightarrow{k_2} E\text{-}O\overset{O}{\overset{\|}{C}}R \xrightarrow[HOH]{k_3} E + RCOO^- + X^-$$

If deacylation is completely rate determining for the hydrolysis of AcF-pNP and sucAAPF-SBzl (that is, $k_{cat,ester} = k_3$), then $K_m = K_s[k_{cat,ester}/(k_2 + k_{cat,ester})]$ and $k_{cat} = k_2k_{cat,ester}/(k_2 + k_{cat,ester})$ (*15*). The values for the Ch reaction with AcF-NH_2 are in agreement with other reports (*15*).

Substrate	K_s (M)	k_2 (s^{-1})	k_3 (s^{-1})
	Ch		
AcF-NH_2	0.023	0.43	60*
sucAAPF-AMC	5.2×10^{-5}	160	52
sucAAPF-pNA	1.5×10^{-3}	850	52
	Tr		
AcF-NH_2	1.0×10^{-3}	0.00022	30*
sucAAPF-AMC	5.8×10^{-4}	0.0066	36
sucAAPF-pNA			36
	D189S		
AcF-NH_2	0.17	0.018	39*
sucAAPF-AMC	7.8×10^{-4}	0.058	33
sucAAPF-pNA	0.015	0.29	33
	Tr→Ch[S1+L1+L2]		
AcF-NH_2	0.065	0.018	33*
sucAAPF-AMC	≥0.014	≥67	37
sucAAPF-pNA	0.011	20	37

*Enzymatic hydrolysis of AcF-pNP was measured at pH 6.5 to avoid nonenzymatic hydrolysis. Hydrolysis of sucAAPF-SBzl at pH 6.5 differed slightly from that at pH 8.0 for all four enzymes, indicating that in this range there is little pH dependence for k_3. Therefore, k_{cat} for AcF-pNP hydrolysis at pH 6.5 is a reasonable estimate for k_3 of AcF-NH_2 hydrolysis at pH 8.0.

than Tr. This low degree of specificity reflects the discrimination of substrate binding: these substrates are so reactive that any binding likely results in acylation. In contrast, there is a 10^5-fold differential in the activity of Tr and Ch on sucAAPF-AMC; 10^4-fold is a result of discrimination at the acylation step (k_2).

The mutant Tr→Ch[S1+L1+L2] further illustrates the relative importance of binding in substrate specificity. This enzyme hydrolyzes sucAAPF-AMC much better than either Tr or D189S, although both Tr and D189S bind sucAAPF-AMC tenfold better than does Tr→Ch[S1+L1+L2]. Substrate binding in Tr→Ch[S1+L1+L2] is severely impaired as compared to Ch. Thus, the mutant lacks a ground-state enzyme-substrate interaction that can be overcome with high concentrations of substrate. Perhaps Tr→Ch[S1+L1+L2] has a "Tr-like" conformation but can be driven into a "Ch-like" conformation with excess substrate. These results imply the existence of additional, unrecognized, structural determinants of substrate specificity.

For Ch, the acylation rate constant k_2 is 300- to 2000-fold greater for the oligopeptide amide substrates than for AcF-NH_2; K_s is decreased only 10- to 200-fold. Therefore, occupation of the subsites of Ch by an oligopeptide amide increases acylation as well as substrate binding (*20*). This acceleration of acylation is not observed for Tr or D189S with sucAAPF-AMC or sucAAPF-pNA substrates. Chymotrypsin, but not Tr or D189S, has the ability to translate binding of these extended substrates into acceleration of acylation. Similarly, Tr→Ch [S1+L1+L2], like Ch, can translate binding in the extended binding subsites into a 10^3-fold increase in acylation of sucAAPF-AMC or sucAAPF-pNA relative to AcF-NH_2. The k_2 and k_3 rate constants for Tr→Ch[S1+L1+L2]–catalyzed hydrolysis of these oligopeptide amides are comparable to those for the Ch reaction. Thus, loops 1 and 2 somehow stabilize the extended transition state of oligopeptide amide hydrolysis and mediate the acceleration of acylation by the P2-P4 residues.

The complex of Tr with bovine pancreatic trypsin inhibitor can be used to infer the subsite contacts between Tr and oligopeptide substrates (Fig. 1B). The P2 through P4 inhibitor residues form an antiparallel β sheet with enzyme residues 214 through 218. These residues also form a wall of the S1 binding pocket. These interactions provide a means to translate the binding energy of the P2 through P4 substrate residues into catalysis. The binding energy might be used to rotate the amide C–N bond, thereby breaking the resonance stabilization and distorting the molecule toward the transition state. Such a strain mechanism would not have a dramatic effect on ester hydrolysis because the ester C–O bond rotates more freely than an amide C–N bond. Many esterases and lipases have catalytic triads similar to those found in serine proteases (*21*). From our results, we predict that these esterases and lipases have little amidase activity (*22*). Further, their conversion to amidases by mutagenesis may be difficult indeed.

The walls of the binding pockets of Ch and Tr can be aligned with a root-mean-square deviation of less than 0.6 Å (Fig. 1), which is comparable to the deviation observed between two independent Ch structures (*5, 23*). However, differences of as little as 0.01 Å in critical bond lengths can have profound effects on chemical reactivity (*24*). Loops 1 and 2 are not structural components of the S1 or the extended binding sites. Therefore, either loops 1 and 2 produce an undetectable change in tertiary structure or they provide the structural flexibility required by the reaction dynamics. These and further mutant enzymes may illuminate this fundamental mystery of serine protease catalysis.

REFERENCES AND NOTES

1. C. deHaen, H. Neurath, D. C. Teller, *J. Mol. Biol.* **92**, 225 (1975).
2. T. A. Steitz, R. Henderson, D. M. Blow, *ibid.* **46**, 337 (1969).
3. R. M. Stroud, L. M. Kay, R. E. Dickerson, *ibid.* **83**, 185 (1974).
4. The nomenclature (P*n*, . . . , P2, P1, P1′, P2′, . . . , P*n*′) denotes the substrate amino acid residues, where P1-P1′ is the hydrolyzed bond. (*Sn*, . . . , S2, S1, S1′, S2′, . . . , *Sn*′) are the corresponding enzyme binding sites [I. Schechter and A. Berger, *Biochem. Biophys. Res. Commun.* **27**, 157 (1968)].
5. G. H. Cohen, E. W. Silverton, D. R. Davies, *J. Mol. Biol.* **148**, 449 (1981).
6. R. Huber and W. Bode, *Acc. Chem. Res.* **11**, 114 (1978).
7. A third pancreatic serine protease, elastase, cleaves after small aliphatic residues; in this enzyme the entrance to the S1 binding pocket is occluded by Val^{216} and Thr^{226} (both residues are Gly in Ch and Tr) [D. M. Shotton and H. C. Watson, *Nature* **225**, 811 (1970)].
8. L. Gráf *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 4961 (1988).
9. Mutant nomenclature: D189S, rat Tr II with Asp^{189} replaced by Ser; Q192M, Tr with residue Gln^{192} replaced by Met; I138T, Tr with Ile^{138} replaced by Thr; and T219, Tr with Thr^{219} inserted (Fig. 2).
10. Specific ester catalysis by Ch is diffusion-controlled [A. C. Brouwer and J. F. Kirsch, *Biochemistry* **21**, 1302 (1982)].
11. R. C. Thompson and E. R. Blout, *ibid.* **12**, 57 (1973); C.-A. Bauer, R. C. Thompson, E. R. Blout, *ibid.* **15**, 1291 (1976); C.-A. Bauer, *ibid.* **17**, 375 (1978); R. L. Stein, A. M. Strimpler, H. Hori, J. C. Powers, *ibid.* **26**, 1301 (1987).
12. The mutant Tr→Ch[S1−1] was constructed from wild-type template with the mutagenic oligonucleotides 5′-GAG-GGA-GGC-AAG-TCT-TCC-TGC-CAG-GGT-3′ (D189S mutation), 5′-T-TCC-TGC-ATG-GGT-GAC-TCT-GGT-GG-3′ (Q192M mutation), and 5′-CT-CAG-TGC-CTG-ACG-TCT-GGC-TGG-G-3′ (I138T mutation) [T. A. Kunkel, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 488 (1985)]. Tr→Ch[S1−2] and Tr→Ch[S1−3] were constructed from template D189S and Tr→Ch[S1−1], respectively, with the mutagenic oligo 5′-TGG-GGC-TAT-GGC-ACT-TGT-GCC-CTG-CCA-3′. Tr→Ch[S1+L1] was constructed from Tr→Ch[S1−3] template and mutagenic oligo 5′-AAC-ATG-GTC-TGT-GTT-GGC-GCC-TCG-GGA-GGC-TCT-TCC-TGC-ATG-GGT-GAC-3′. Tr→Ch[S1+L2] and Tr→Ch[S1+L1+L2] were constructed from Tr→Ch[S1−3] and Tr→Ch[S1+L1] templates, respectively, and mutagenic oligo 5′-TGG-GGC-TCT-GGC-ACT-TGT-TCC-ACT-TCG-ACT-CCT-GGT-GTG-TAC-ACC-3′. The desired mutations were verified with sequencing.
13. In elastases the sequence of these two loops is different from Ch or Tr and is also characteristically conserved (*1*).
14. N. Yoshida *et al.*, *Biochemistry* **19**, 5799 (1980); K. Nakajima, J. C. Powers, R. M. Ashe, M. Zimmerman, *J. Biol. Chem.* **254**, 4027 (1979).
15. B. Zerner, R. P. M. Bond, M. L. Bender, *J. Am. Chem. Soc.* **86**, 3669 (1964).
16. L. Polgar, *Mechanisms of Protease Action* (CRC Press, Boca Raton, FL, 1989), chap. 3.
17. U. Christensen and H. H. Ipsen, *Biochim. Biophys. Acta* **569**, 177 (1979); R. L. Stein *et al.*, *J. Am. Chem. Soc.* **106**, 796 (1984).
18. E. G. DelMar, C. Largman, J. W. Brodrick, M. C. Geokas, *Anal. Biochem.* **99**, 316 (1979); J. W. Harper, G. Ramirez, J. C. Powers, *ibid.* **118**, 382 (1981).
19. If the deacylation step is rate determining, the presence of a nucleophile such as alaninamide will

increase k_{cat} [M. L. Bender, G. E. Clement, C. R. Gunter, F. J. Kezdy, *J. Am. Chem. Soc.* **86**, 3697 (1964)]. Alaninamide increased k_{cat} equally for the hydrolysis of both sucAAPF-SBzl and sucAAPF-AMC by Ch, which indicates that deacylation is rate determining for both compounds. Alaninamide also increased k_{cat} for the hydrolysis of sucAAPF-SBzl by D189S and mutant Tr→Ch[S1+L1+L2], which confirms that deacylation is the rate determining step in these reactions (L. Hedstrom, unpublished experiments).

20. In the elastase reaction, increasing peptide length also increases acylation in preference to binding or deacylation [R. C. Thompson and E. R. Blout, *Proc. Natl. Acad. Sci. U.S.A.* **67**, 1734 (1970)].
21. L. Brady *et al.*, *Nature* **343**, 767 (1990); F. K. Winkler, A. D'Arcy, W. Hunziker, *ibid.*, p. 771; M. M. G. M. Thunnissen *et al.*, *ibid.* **347**, 689 (1990); J. D. Schrag, Y. Li, S. Wu, M. Cygler, *ibid.* **351**, 761 (1991); J.-P. Wery *et al.*, *ibid.* **352**, 79 (1991); J. L. Sussman *et al.*, *Science* **253**, 872 (1991).
22. Acetylcholinesterase hydrolyzes acetylcholine 10^5-fold faster than the analogous amide (k_{cat}/K_m), which is consistent with the relative reactivity of esters and amides [D. E. Moore and G. P. Hess, *Biochemistry* **14**, 2386 (1975)]. In contrast, Ch hydrolyzes sucAAPF-SBzl only sixfold faster than sucAAPF-pNA.
23. R. A. Blevins and A. Tulinsky, *J. Biol. Chem.* **260**, 4264 (1985).
24. P. G. Jones and A. J. Kirby, *J. Chem. Soc. Chem. Commun.* **1979**, 288 (1979).
25. We found that Tr, D189S, and Tr→Ch[S1+L1+L2] were also expressed in yeast culture medium by fusing the trypsinogen coding sequences to the α-factor leader sequence as described for carboxypeptidase A1 [M. A. Phillips, R. Fletterick, W. J. Rutter, *J. Biol. Chem.* **265**, 20692 (1990)]. The final expression vector pYT contains the inducible alcohol dehydrogenase–glyceraldehydephosphate dehydrogenase (ADH-GAPDH) promoter and regulatory regions, the GAPDH transcription terminator, and *amp*, *ura3*, and *leu2d* markers. Medium was isolated from a 2-liter culture by centrifugation and was concentrated on a phenyl-Sepharose column. Trypsinogen was activated and purified (*26*). Typically, 30 mg of purified Tr was obtained. Trypsins isolated from the yeast expression system have kinetic constants identical to those of enzymes isolated from the *Escherichia coli* system.
26. Rat trypsinogen II mutants were expressed in the periplasm of *E. coli* and purified with $(NH_4)_2SO_4$ precipitation and anion exchange chromatography [L. Gráf *et al.*, *Biochemistry* **26**, 2616 (1987)]. The trypsinogens were activated by enterokinase treatment, and Tr was isolated by affinity chromatography on SBTI Sepharose (Sigma Chemical Co.). Typically, 1 mg of purified mutant Tr was obtained. The reported activities are not likely the result of misfolded protein or contaminating activities. First, all of the mutants bound quantitatively to SBTI Sepharose when activated by enterokinase. The mutant trypsins were stable at 4°C for several months. The activity of mutant Tr→Ch[S1+L1+L2] is observed in several independent preparations of Tr→Ch[S1+L1+L2] and not in any other mutant Tr preparation. Mutants D189S and Tr→Ch[S1+L1+L2] were expressed in both *E. coli* and yeast (*25*), with identical steady-state kinetic constants for the hydrolysis of sucAAPF-AMC and benzoyltyrosylethyl ester. The k_{cat} value for ester hydrolysis is equivalent for Ch, Tr, D189S, and Tr→Ch[S1+L1+L2], which indicates that the activity is unlikely to represent contaminating proteases and that the preparations are fully active.
27. M. Zimmerman, B. Ashe, E. C. Yurewicz, G. Patel, *Anal. Biochem.* **78**, 47 (1977).
28. L. Gráf, I. Boldogh, L. Szilagyi, W. J. Rutter, in *Protein-Structure Function*, Z. H. Zaidi, A. Abbasi, M. Smith, Eds. (TWEL Publishers, Karachi, Pakistan, 1990), pp. 49–65.
29. K. Brady and R. H. Abeles, *Biochemistry* **29**, 7608 (1990).
30. W. Bode and P. Schwager, *J. Mol. Biol.* **98**, 693 (1975).
31. R. Huber *et al.*, *ibid.* **89**, 73 (1974).
32. T. E. Ferrin, C. C. Huang, L. E. Jarvis, R. Langridge, *J. Mol. Graph.* **6**, 13 (1988).
33. J. R. Brown and B. S. Hartley, *Biochem. J.* **101**, 214 (1966).
34. S. D. Pinsky, K. S. LaForge, V. Luc, G. Scheele, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 7486 (1983).
35. K. Titani, L. H. Ericsson, H. Neurath, K. A. Walsh, *Biochemistry* **14**, 1358 (1975).
36. C. S. Craik *et al.*, *J. Biol. Chem.* **259**, 14255 (1984).
37. This paper is dedicated to R. H. Abeles on the occasion of his 65th birthday. This work was supported by NIH grant DK21344. We thank J. Cross, M. Phillips, and M. Garabedian for comments on the manuscript. This work is part of an ongoing collaborative project between the W. J. Rutter laboratory at the University of California at San Francisco and the L. Gráf laboratory at Loránd Eötvös University.

23 October 1991; accepted 13 January 1992

Myotonic Dystrophy Mutation: An Unstable CTG Repeat in the 3′ Untranslated Region of the Gene

MANI MAHADEVAN, CATHERINE TSILFIDIS, LUC SABOURIN, GARY SHUTLER, CHRIS AMEMIYA, GERT JANSEN, CATHERINE NEVILLE, MONICA NARANG, JUANA BARCELÓ, KIM O'HOY, SUZANNE LEBLOND, JANE EARLE-MACDONALD, PIETER J. DE JONG, BÉ WIERINGA, ROBERT G. KORNELUK*

Myotonic dystrophy (DM) is the most common inherited neuromuscular disease in adults, with a global incidence of 1 in 8000 individuals. DM is an autosomal dominant, multisystemic disorder characterized primarily by myotonia and progressive muscle weakness. Genomic and complementary DNA probes that map to a 10-kilobase Eco RI genomic fragment from human chromosome 19q13.3 have been used to detect a variable length polymorphism in individuals with DM. Increases in the size of the allele in patients with DM are now shown to be due to an increased number of trinucleotide CTG repeats in the 3′ untranslated region of a DM candidate gene. An increase in the severity of the disease in successive generations (genetic anticipation) is accompanied by an increase in the number of trinucleotide repeats. Nearly all cases of DM (98 percent or 253 of 258 individuals) displayed expansion of the CTG repeat region. These results suggest that DM is primarily caused by mutations that generate an amplification of a specific CTG repeat.

MYOTONIC DYSTROPHY (DM) IS an autosomal dominant disorder characterized primarily by myotonia and progressive muscle weakness, although central nervous system, cardiovascular, and ocular manifestations also frequently occur (*1*). The disease shows a marked variability in expression, ranging from a severe congenital form that is frequently fatal just after birth to an asymptomatic condition associated with normal longevity. Furthermore, the progression of DM in affected families may exhibit genetic anticipation, an increase in the severity of the disease in successive generations. (*2*).

We have cloned the essential region of human chromosome 19q13.3 that contains the DM locus (*3–5*). We isolated genomic and cDNA probes that map to this region and that detect an unstable, 10-kb Eco RI genomic fragment in individuals affected with DM. This variable length polymorphism is similar to the instability at the fragile X locus (*6–8*). The physical map location and genetic characteristics of the variable length polymorphism are consistent with it being the cause of DM. We now further characterize the variable length polymorphism segment to determine the molecular basis of the unstable genomic region.

We identified, using Southern (DNA) blot analysis, two different DNA polymorphisms mapping within the 10-kb Eco RI genomic fragment (Fig. 1). The polymorphisms are detected by two genomic probes (pGB2.6 and pGP1.5) that map to this Eco RI fragment. In normal individuals, these probes detect an 8.5- or 9.5-kb Hind III or a 9.0- or 10.0-kb Eco RI insertion polymorphism. This variation is due to a 1-kb insertion within a 150-bp Pst I–Bgl I fragment (Fig. 1). The second DNA polymorphism maps to a 1.5-kb Bam HI fragment located 4.0 kb centromeric to the insertion poly-

M. Mahadevan, C. Tsilfidis, L. Sabourin, G. Shutler, C. Neville, M. Narang, J. Barceló, K. O'Hoy, Department of Microbiology and Immunology, University of Ottawa, Ottawa, Ontario, Canada K1H 8L1.
C. Amemiya, and P. J. de Jong, Lawrence Livermore National Laboratories, Livermore, CA 94550.
G. Jansen, Department of Human Genetics, University of Nijmegen, 6500 HB, Nijmegen, the Netherlands.
S. Leblond and J. Earle-Macdonald, Research Institute, Children's Hospital of Eastern Ontario, Ottawa, Canada K1H 8L1.
B. Wieringa, Department of Cell Biology and Histology, University of Nijmegen, 6500 HB, Nijmegen, the Netherlands.
R. G. Korneluk, Department of Microbiology and Immunology, University of Ottawa, Ottawa, Ontario, Canada K1H 8L1 and Research Institute, Children's Hospital of Eastern Ontario, Ottawa, Ontario, Canada K1H 8L1.

*To whom correspondence should be addressed.