Protein and Peptide Letters, Vol. 4, No. 5, pp. 335-342, 1997
Bentham Science Publishers B.V.
0929-8665/94 $ 5.00 + .00

INTRODUCTION OF TRYPSIN SPECIFICITY INTO CHYMOTRYPSIN INHIBITOR II

Zulfiqar Hasan and Robin J. Leatherbarrow*

Department of Chemistry, Imperial College of Science, Technology and Medicine, South Kensington, London SW7 2AY, UK

A unique cysteine residue was introduced into the P_1 position of chymotrypsin inhibitor II (CI2) by oligonucleotide-directed mutagenesis ($Met^{59} \rightarrow Cys$). Chemical modification of CI2M59C with 2-bromoethylamine, 3-bromopropylamine and 2-mercaptoethylamine resulted in the introduction of lysine-sulfur analogues at the P_1 position. Unlike the original CI2, all three modified inhibitors exhibited moderate but effective inhibition of trypsin. The results demonstrate that genetic and chemical modification can be successfully combined to tailor inhibitory specificity.

INTRODUCTION

The combination of genetic engineering and chemical modification provides a way to introduce precise alterations in order to investigate structure / function relationships within proteins. The most utilized and successful approach is cysteine engineering [1-3], which is useful for chemical labeling of proteins and allows the generation of modified proteins containing non-natural amino acid analogues. In this technique, a cysteine is introduced by recombinant DNA technology and subsequently reacted with specific modifying reagents. This overcomes the natural limitation of oligonucleotide-directed mutagenesis to substitution by only 19 other naturally occurring amino acids.

The specificity of trypsin is largely influenced by a negatively charged residue (Asp 189) located at the base of the S_1 specificity pocket (nomenclature of protease subsites from Schechter and Berger [4]). Consequently, trypsin exhibits a narrow substrate specificity profile, with a preference for positively charged groups, such as Lys or Arg, at the P_1 position of substrates or inhibitors. The protease inhibitor CI2, which has a Met residue at P_1, shows no inhibition of trypsin, and has specificity for chymotrypsin and subtilisin [5, 6]. CI2 is a small protease inhibitor found in seeds of barley [7], and like many plant protease inhibitors is thought to have a defensive role against wounding or attack. It is a member of the potato I family of serine protease inhibitors,

and inhibition is mediated by interaction with an exposed loop that is fixed in the characteristic "canonical" conformation found in many different serine protease inhibitor families [8]. This paper reports the tailoring of CI2 specificity towards trypsin, emphasizing the importance of the P_1 residue as the primary determinant of inhibitory specificity in serine proteases [5, 9-12]. Oligonucleotide-directed mutagenesis was carried out to substitute the P_1 methionine by a cysteine residue. The cysteine residue in the mutant CI2M59C was then chemically elaborated to *S*-(β-aminoethyl)cysteine, *S*-(β-aminopropyl)cysteine and *S*-(β-aminoethylthio)cysteine by reaction with 2-bromoethylamine, 3-bromopropylamine and 2-mercaptoethylamine, respectively. The chemically modified mutant inhibitors thus contained positively charged sulfur-containing analogues of lysine with different chain lengths. These variants were then assayed for specificity towards trypsin and their inhibitory potencies were measured.

MATERIALS AND METHODS

Site-directed mutagenesis, expression and purification. The mutant gene CI2M59C was constructed by the Kunkel method of mutagenesis [13] using the mutagenic primer 5′-CGATATTCGCAGGTCACA-3′, where the nucleotides underlined direct the mutation Met^{59}(ATG) → Cys (TGC). Mutagenesis was performed on the CI2 gene sub-cloned from pPO1 [14] into M13mp18. Mutant clones were identified by differential hybridization [15] and the mutation confirmed by DNA sequencing [16]. The CI2M59C gene was further sub-cloned into the plasmid pT7-7 [17], and mutant protein expressed from this construct in *E. coli* K38 cells [18] harboring the plasmid pGP1-2 [19]. The protein was purified according to the published protocol for wild-type CI2 [5] and was generally greater than 95% pure, as judged by SDS-PAGE.

Titration of the sulfhydryl content with 5,5′-dithiobis 2-nitrobenzoic acid (DTNB). Titration of the sulfhydryl content of CI2M59C was carried out by the method of Ellman [20].

Preparation of CI2M59S-(β-aminoethyl)cysteine, CI2M59S-(β-aminopropyl)cysteine and CI2M59S-(β-aminoethylthio)cysteine. Chemical modification reactions were carried out after dissolving 10 mg of CI2M59C protein in 1 mL 0.1 M Tris-HCl (pH 8.2) containing 10 mM DTT and incubating overnight to reduce the inhibitor to a completely monomeric conformation. 2-bromoethylamine, 2-bromopropylamine or 2-mercaptoethylamine (Sigma Chemical Co.) was then added to a final concentration of 50 mM, nitrogen gas bubbled through the solution to purge air from the system and the reaction mixture incubated in the dark at 37°C. Aliquots (3 μl) were removed at various time intervals during the reaction and resolved by isoelectric focusing (IEF) using the Phast™ system (Pharmacia Biotech) to monitor the progress of the reaction. On termination of the reaction, excess modifying and reducing agents were removed by desalting on a 10 mL

Sephadex G-15 gel filtration column. The protein was then dialyzed exhaustively against H_2O and subsequently lyophilized. In each case, the modified inhibitor was further purified by anhydrotrypsin affinity chromatography.

Anhydrotrypsin affinity chromatography. The chemically modified inhibitors were purified by anhydrotrypsin affinity chromatography. Anhydrotrypsin was prepared as described previously [21] and coupled to CNBr-activated Sepharose® 4B (Pharmacia Biotech) according to the manufacturers instructions. A one-step purification procedure separated the modified inhibitors from unmodified protein, thus allowing the recovery of homogeneous preparations of the aminoalkylated inhibitors. The modified inhibitor preparations were dissolved in buffer A (0.1 M Tris-HCl pH 8.0, 0.1 M NaCl, 0.12 M $CaCl_2$) and loaded onto the anhydrotrypsin affinity column at a flow rate of 1 mL min^{-1}. Under these conditions the aminoalkylated inhibitors bound to the column. Bound protein was eluted with buffer B (0.01 M HCl, 0.1 M KCl), dialyzed exhaustively against H_2O and subsequently lyophilized. Modification of the protein was verified by electrospray mass spectrometry and purity confirmed by IEF.

Kinetic analysis. Bovine pancreatic trypsin (Sigma) was assayed in 0.144 M Tris-HCl (pH 7.78) at 25°C with L-benzoylarginine-*p*-nitroanilide (L-BAPNA, Sigma) as substrate [22]. Trypsin concentrations were determined by active site titration using *p*-nitrophenyl-*p*-guanidinobenzoate (NPGB, Sigma) as described by Chase and Shaw [23]. Subsequently, values of k_{cat} and K_m (2.4 s^{-1} and 2.16 mM respectively) for trypsin with L-BAPNA were used to calculate enzyme concentrations from the initial rates of substrate hydrolysis. The K_i value of each mutant inhibitor was determined by the equilibrium method [24]. Data points sampled were analyzed by non-linear regression using the *GraFit* computer program [25]. The apparent equilibrium dissociation constant K_i' (K_i in the presence of substrate) was determined by fitting data to Equation 1 [26].

$$v_i = v_o (K_i' \quad \{K_i' + [I_o]\}) \tag{1}$$

In this equation, v_i and v_o are the steady-state inhibited and uninhibited rates, respectively, and $[I_o]$ is the total inhibitor concentration. The true K_i value (at zero substrate) was calculated from Equation 2.

$$K_i = K_i' / (1 + [S] / K_m) \tag{2}$$

The rate of inhibitor hydrolysis by trypsin was determined by measuring recovery of the enzyme activity of an inhibited reaction over time. Sufficient inhibitor was added to give approximately 50% inhibition. Initial (steady-state) rates were then plotted versus time of incubation and the data fitted to a first order rate equation (Equation 3).

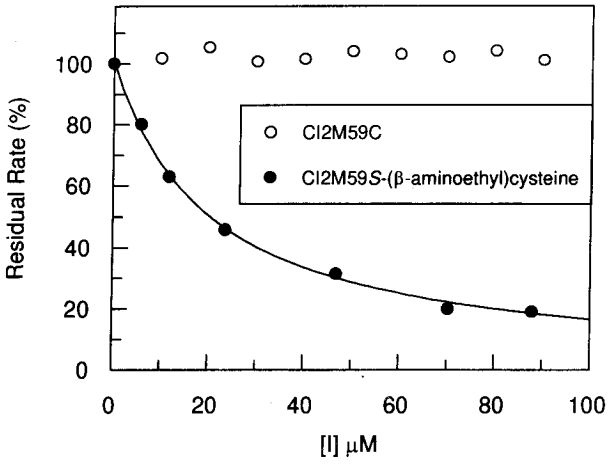
$$v = (v_\infty - v_o)(1 - e^{-k\,t}) + v_o \qquad (3)$$

In this equation, the rate v varies with time t, k represents the first order rate constant for hydrolysis of the inhibitor, and v_∞ and v_o are the final and starting rates respectively.

RESULTS

The absence of cysteine residues in wild-type CI2 allowed a convenient way of combining genetic and chemical modification to produce the inhibitor variants. The Kunkel method of mutagenesis [13] was successfully used to construct the mutant gene CI2M59C. The mutant gene was then transferred to the expression vector pT7-7 and high level expression achieved. As CI2M59C carries an exposed thiol group, intermolecular disulfide bonds formed, resulting in inactive inhibitor dimers (data not shown). Inhibitory activity against subtilisin BPN′ was restored upon reduction of the protein with DTT, and the CI2M59C protein remained a potent inhibitor of subtilisin BPN′ (K_i value of 2.1×10^{-11} M; wild-type protein has K_i value of 2.9×10^{-12} M [5]). Thereafter, the mutant protein was modified following its complete reduction.

Figure 1 Inhibition of trypsin by CI2M59*S*-(β-aminoethyl)cysteine. Enzyme and inhibitor were pre-incubated for 15 min prior to measuring the rate of substrate hydrolysis, to allow time for EI complex formation. Enzyme and substrate concentrations were 10 nM and 0.5 mM, respectively. Steady-state rates are expressed as a percentage of the uninhibited rate. The apparent equilibrium dissociation constant K_i' was determined by non-linear fitting of the data to Equation 1. CI2M59C was used as a negative control.



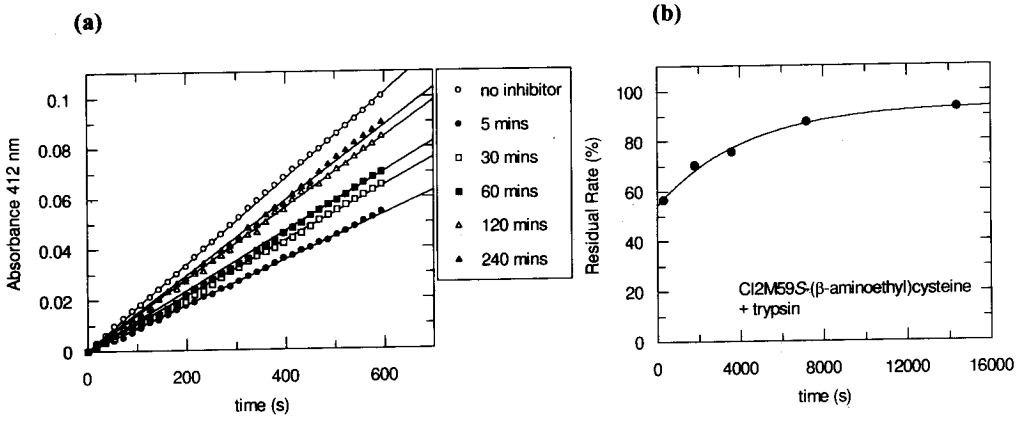
As modification of CI2M59C with 2-bromoethylamine, 3-bromopropylamine and 2-mercaptoethylamine introduced a positive charge on the protein, the progress of the modification reactions could be monitored by IEF. For each of the three modifications, the extra positive charge caused a shift in pI of the protein from 6.2 to 8.3. The specificity of the 3 reagents for cysteine residues was confirmed by performing identical modification reactions with wild type CI2 (data not shown). As expected, no shift in pI was observed on reaction of wild type CI2 with any of these reagents. The progress of modification reactions was monitored every hour for up to 3 hours, but the reaction did not reach completion (the extent of modification after 3 hours was approximately 80-90%). Incubation for longer periods did not increase the modification efficiency (data not shown). Consequently, the modified inhibitors were purified by anhydrotrypsin affinity chromatography to remove unmodified protein and their purity subsequently verified by IEF. Electrospray mass spectroscopy confirmed that the modified proteins had the expected molecular weights, and so carried a single modification.

Kinetic characterization revealed all 3 modified inhibitors to be inhibitors of trypsin. A representative set of data is shown in Figure 1. Of the modifications carried out, reaction with 2-mercaptoethylamine produced the most potent inhibitor, with a K_i value of 7.4×10^{-6} M (Table 1). Reaction with 2-bromoethylamine and 3-bromopropylamine produced trypsin inhibitors with higher K_i values as listed in Table 1. As both wild type CI2 and CI2M59C were observed to undergo hydrolysis by trypsin over extended periods of incubation, the three aminoalkylated inhibitors were similarly investigated. Hydrolysis was studied both qualitatively (by SDS-PAGE, data not shown) and quantitatively (by determining the first order rate constant for hydrolysis). Representative data showing the hydrolysis of CI2M59*S*-(β-aminoethyl)cysteine by trypsin is presented in Figure 2. Hydrolysis rates for CI2M59*S*-(β-aminoethyl)cysteine, CI2M59*S*-(β-aminopropyl)cysteine and CI2M59*S*-(β-aminoethylthio)cysteine are listed in Table 1. Hydrolysis rates could not be determined for wild type CI2 or CI2M59C as these proteins do not inhibit trypsin and thus no change in the rate of substrate hydrolysis could be observed.

Table 1 Kinetic constants for inhibition of trypsin by CI2M59*S*-(β-aminoethyl)cysteine, CI2M59*S*-(β-aminopropyl)cysteine and CI2M59*S*-(β-aminoethylthio)cysteine and hydrolysis rates of the modified inhibitors. Standard errors from the data fitting are in each case < 10%.

Inhibitor	P_1 side chain	K_i (μM)[a]	Hydrolysis rate (mole / mole / sec) $\times 10^{-4}$ [b]	$t_{1/2}$ [c] (min)
wild type CI2	$-CH_2CH_2SCH_3$	no inhibition	N.D[d]	N.D
CI2M59C	$-CH_2SH$	no inhibition	N.D[d]	N.D
CI2M59*S*-(β-aminoethyl) cysteine	$-CH_2SCH_2\ CH_2NH_3^+$	11.6	2.3	50.2
CI2M59*S*-(β-aminopropyl)cysteine	$-CH_2SCH_2CH_2CH_2NH_3^+$	12.5	2.1	55.0
CI2M59*S*-(β-aminoethylthio)cysteine	$-CH_2S-SCH_2CH_2NH_3^+$	7.4	8.7	13.3

Figure 2 Hydrolysis of CI2M59*S*-(β-aminoethyl)cysteine by trypsin. (a) Progress curves showing the increase in enzymic rate of an inhibited reaction over time. The enzyme and inhibitor concentrations were 10 nM and 20 μM, respectively. Enzyme and inhibitor were incubated at 25°C for various time intervals (5, 30, 60, 120 and 240 min). The chromogenic substrate, L-BAPNA, was then added to each cuvette to a final concentration of 0.5 mM, and the change in absorbance at 412 nm recorded over 30 min. Initial (steady-state) rates were determined by linear fitting of the data. (b) Time course of hydrolysis. Initial rates calculated from (a) above, are expressed as a percentage of the uninhibited rate. The data obtained were fitted to the Equation 3 by non-linear regression to determine the first order rate constant for hydrolysis of the modified inhibitor.



(a) K_i values were experimentally determined using the equilibrium method [20].

(b) Hydrolysis rates (mole of inhibitor hydrolyzed / mole of enzyme / second) were experimentally determined. Data were fitted to equation 3 by non-linear regression. Each value listed represents the mean value from 3 independent determinations.

(c) The half-life for inhibitor hydrolysis is given by Equation 4.

$$t_{1/2} = \frac{\ln 2}{k} \qquad (4)$$

Values for $t_{1/2}$ were obtained by substituting experimentally determined values for k (hydrolysis rate) into Equation 3.

(d) Slow hydrolysis and lack of inhibition prevented quantitative calculation of hydrolysis rates.

DISCUSSION

Previous work has shown that aminoethylation of a mutant α_1-proteinase inhibitor having a cysteine residue introduced at P_1 can lead to redirection of specificity towards trypsin [27]. In the present study, chemical modification of CI2M59C with 2-bromoethylamine, 3-bromopropylamine and 2-mercaptoethylamine was successful in creating 3 trypsin inhibitors with K_i values in the μM range. The variant inhibitors differed from each other in the length of their P_1 side chain, although no clear trend was evident between P_1 side chain lengths and K_i values. However, it is interesting to note that all three inhibitor variants were approximately 1000-fold less potent than the mutant CI2M59K (5.6 nM) reported previously [5]. This difference is most surprising for CI2M59*S*-(β-aminoethyl)cysteine as the modified P_1 residue is isosteric with lysine, differing only in the

substitution of a methylene (-CH_2-) linkage by a -S- linkage. However, one factor that needs to be considered is the observation that all 3 aminoalkylated inhibitors are rapidly hydrolyzed by trypsin. Consequently for CI2M59*S*-(β-aminoethylthio)cysteine ($t_{1/2}$ = 13.3 min), in the equilibrium determination of K_i values, true steady-state inhibited rates were almost certainly not measured due to its rapid hydrolysis during the time course of the experiment. This means that the true K_i values determined for this modified inhibitor may be lower than the values in Table 1 suggest. It has also been reported that the mutant CI2M59K was cleaved by trypsin, but only after prolonged (up to 16 hours) incubation [5]. In this case, an accurate value for K_i could be determined as cleavage appeared to take place slowly whilst the establishment of equilibrium occurred relatively quickly. An alternative explanation that may partially account for the higher than expected K_i values of the modified inhibitors, is the presence of -S- linkages within the P_1 side chains. Although sulfur has an atomic radius similar to that of -CH_2-, it contains unpaired electrons and these may interact unfavorably within the environment of the S_1 pocket. Also, the bond angle between -C-C- (109.5°) is not the same as in aliphatic sulfides (-C-S-C-, 105°) or disulfides (-S-S-, 103°-107°) [28]. As a result, the terminal -NH_3^+ groups of the P_1 side chain may not be positioned optimally for interaction with Asp 189 at the base of the S_1 pocket. This unfavorable geometry may explain the high K_i values of the modified inhibitors and the consequent rapid hydrolysis. Thus, mutation of the P_1 residue resulted in altered inhibitory activity towards trypsin and owing to less potent inhibition as compared to CI2M59K, the inhibitor variants acted as a better substrate.

In summary, we have studied the effect of introducing lysine-like residues at the P_1 site of CI2 by cysteine engineering. Additionally, the effect of the side chain length of the P_1 residue was also investigated by chemical modification to analogues of increasing chain lengths. CI2 was successfully converted to an inhibitor of trypsin but the analogues suffered more rapid hydrolysis by trypsin than is found for natural trypsin-specific inhibitors.

ACKNOWLEDGEMENTS

We would like to thank Professor H. Morris for performing the electrospray mass spectroscopy.

REFERENCES

[1] Kim, D.W., Yoshimura, T., Esaki, N., Satoh, E. and Soda, K. (1994) *J. Biochem* **115,** 93-97.
[2] Gilardi, G., Zhou, Q., Hibbert, L. and Cass, A.E.G. (1994) *Anal. Chem.* **66,** 3840-3847.
[3] Profy, A.T. and Schimmel, P. (1988) *Prog. Nucleic Acid Res. and Mol. Biol.* **35,** 1-26.
[4] Schechter, I. and Berger A. (1967) *Biochem. Biophys. Res. Commun.* **27,** 157-162.
[5] Longstaff, C.A., Campbell, A.F. and Fersht, A.R. (1990) *Biochemistry* **29,** 7339-7347.
[6] Hasan, Z. (1996) Ph.D. Thesis University of London.
[7] Boisen, S., Andersen, C.Y. and Hejgaard, J. (1981) *Physiol. Plant.* **52,** 167-176.
[8] Bode, W. and Huber, R. (1992) *Eur. J. Biochem.* **204,** 433-451.
[9] Jallat, S., Carvallo, D., Tessier, L.H., Roeklin, D., Roitsch, C., Ogushi, F., Crystal, R.G. and Courtney, M. (1986) *Protein Eng* **1,** 29-35.

[10] Takeuchi, Y., Noguchi, S., Yoshinori, S., Kojima, S., Kumagai, I., Kin-ichiro, M., Nakamura, K.T. and Mitsui, Y. (1991) *Protein Eng* **4,** 501-508.
[11] Bjork, I., Ylinenjarvi, K., Olson, S.T. and Bock, P.E. (1992) *J. Biol. Chem.* **267,** 1976-1982.
[12] Cooper, S.T. and Church, F.C.(1995) *Biochim. Biophys. Acta.* **1246,** 29-33.
[13] Kunkel, T.A. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 488-492.
[14] Osmark, P., Sorensen, P. and Poulsen, F.M. (1993) *Biochemistry* **32,** 11007-11015.
[15] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) "Molecular Cloning: A laboratory Manual", 2nd Edition, Cold Spring Harbour Laboratory Press, New York.
[16] Sanger, F., Niklen, S. and Coulsen, A.R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463-5467.
[17] Tabor, S (1990) "Current Protocols in Molecular Biology", Asubel, F.A., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. (Eds.), Greene Publishing and Wiley-Interscience, New York.
[18] Russel, M. and Model, P. (1984) *J. Bacteriol.* **159,** 1034-1039.
[19] Tabor, S. and Richardson, C.C. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 1074-1078.
[20] Ellman, G.L. (1959) *Arch. Biochem. Biophys.* **82,** 70-77.
[21] Ako, H., Ryan., C.A. and Foster, R.J. (1972) *Biochem. Biophys. Res. Commun.* **46,** 1639-16.
[22] Erlanger, B.F., Kokowsky, N. and Cohen, W. (1961) *Arch. Biochem. Biophys.* **95,** 271-278.
[23] Chase, T. and Shaw, E. (1967) *Biochem. Biophys. Res. Commun.* **29,** 508-514.
[24] Tonomura, B., Senda, M., Tsura, D., Komiyama, T., Miwa, M., Akaska, K., Kainosho, M., Tsuji, T., Takahashi, K., Hiromi, K., Ikenaka, T. and Murao, S. (1985) "Protein Protease Inhibitor—The Case of *Streptomyces* Subtilisin Inhibitor", Hiromi, K., Akaska, K., Mitsui, Y., Tonomura, B. and Murao, S. (Eds.), Elsevier Science Publishers B.V. (Biomedical Division), 291-362.
[25] Leatherbarrow, R.J. (1992) *GraFit* Version 3.0, Erithacus Software Ltd., Staines, U.K.
[26] Hejgaard, J., Dam, J., Petersen, L.C. and Bjorn, S.E. (1994) *Biochim. Biophys. Acta.* **1204,** 68-74.
[27] Matheson, N.R., Gibson, H.L., Hallewell, R.A., Barr, P.J. and Travis, J. (1986) *J. Biol. Chem.* **261,** 10404-10409.
[28] Sykes, P. (1981) "A Guidebook to Mechanism in Organic Chemistry" 5th ed. Longman, London

Received on May 15, 1997, accepted on September 5, 1997