Protein Engineering vol.10 no.9 pp.991–997, 1997

Mutagenesis, biochemical characterization and X-ray structural analysis of point mutants of bovine chymosin

Mark G.Williams[3], Julie Wilsher, Phil Nugent, Alan Mills, V.Dhanaraj[3], Milan Fabry[1], Juraj Sedlacek[1], J.M.Uusitalo[2], M.E.Penttila[2], Jim E.Pitts and Tom L.Blundell[3,4]

Laboratory of Molecular Biology, Department of Crystallography, Birkbeck College, University of London, Malet Street, London WC1E 7HX, UK, [1]Department of Gene Manipulation, Institute of Molecular Genetics, Academy of Sciences of the Czech Republic, Flemingovo nam. 2, 166 37 Prague 6, Czech Republic and [2]The Biotechnical Laboratory, VTT, Espoo, Finland

[3]Present address: Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QW, UK
[4]To whom correspondence should be addressed

Chymosin B point mutants, A115T and G243D (chymosin A), were expressed in *Escherichia coli and Trichoderma reesei* respectively, characterized biochemically, crystallized and studied by X-ray analysis at 2.3 and 2.8 Å resolutions respectively. The three-dimensional structures showed that the mutations gave rise to local conformational changes only when compared with that of chymosin B. Kinetic analysis of the A115T mutant with a six residue synthetic peptide revealed a reduction in K_m with respect to the wild type, possibly caused by the small local changes in the vicinity of S1 and S3. Although, kinetic analyses of the G243D mutant using the short substrate showed reduced catalytic activity, use of a 15 residue substrate based on residues 98–112 of κ-casein, the natural substrate, revealed an increase in the k_{cat} compared with chymosin B, probably a consequence of the charge introduced that may interact with the substrate between P4 and P8.

Keywords: aspartic proteinase/chymosin/mutant/substrate/crystal structure

Introduction

Aspartic proteinases have been traditionally exploited in food processing including cheese using chymosin from calf stomach (Foltmann, 1966; 1970) and plant enzymes (Brodelius *et al.*, 1993; Cordeiro *et al.*, 1994), soya using mould enzymes such as rhizomucor pepsin (Newman *et al.*, 1993) and cocoa which exploits a natural aspartic proteinase found in the cocoa bean (Voigt *et al.*, 1995). These processes exploit the relatively specific proteolysis at mildly acidic conditions shown by these enzymes in contrast to the much broader specificity of bovine pepsin in the very acidic environment of the stomach. The quality of the food product is determined by the extent of proteolysis, which gives rise not only to clotting but also to small peptides that impart particular tastes.

The specificity of aspartic proteinases depends partly on the degree of denaturation at the pH of the process and partly on the binding pockets along the long active site cleft (Davies, 1990). X-Ray structural and kinetic studies have described a series of specificity pockets, S5 to S4′ (nomenclature of Schechter and Berger, 1967), on either side of the catalytic aspartates; these vary in size, hydrophobicity, electrostatic potential and hydrogen bonding capability both with respect to each other and between different enzymes (Scarborough *et al.*, 1993; Scarborough and Dunn, 1994; Lowther *et al.*, 1995; Rao-Naik *et al.*, 1995; Gustchina *et al.*, 1996).

The early studies of cleavage of a range of polypeptide substrates by bovine pepsin showed that in general the preferred cleavage site is between two hydrophobic residues. Additionally, the preference for particular residues and the degree of hydrophobicity falls off on either side of the scissile bond. X-Ray structural studies of aspartic proteinases with a series of peptide-like inhibitors (Bott *et al.*, 1982; James *et al.*, 1982; Foundling *et al.*, 1987; Dhanaraj *et al.*, 1992; James *et al.*, 1992; Suguna *et al.*, 1992) suggest that the substrate has an extended conformation, making hydrogen bonds with the protein in both parallel and anti-parallel β-sheet interactions on either side of the active site, with a weak pseudo-symmetry relating the residues on either side of the scissile bond. Thus, residues at P1 and P1′ are located in deep hydrophobic pockets that can accommodate a large aliphatic group (e.g. leucyl or cyclohexanylalanyl) or an aromatic sidechain such as phenylalanyl. Residues at P2 and P2′ occupy pockets which are still quite hydrophobic. The inhibitor studies indicate that the conformation of bound substrates at P3′ probably depends very much on the enzyme and the substrate sequence, whereas it is more tightly determined at P3 by specific hydrogen bonds and the specificity pocket S3 which can be quite hydrophobic in enzymes such as renin.

Site-directed mutagenesis studies of Dunn, Kay and co-workers (Scarborough and Dunn, 1994; Rao-Naik *et al.*, 1995) have sought to vary systematically the sequence of peptide substrates and also to vary the specificity subsites using site-directed mutagenesis. For example, mutations of the Met287 residue in the S2 subsite in human cathepsin D showed that substitution by more polar amino acids produced active mutant enzymes with significantly altered substrate specificity. Similar mutagenesis studies by Beppu and co-workers, first on chymosin (Suzuki *et al.*, 1989) and more recently rhizomucor pepsin (Park *et al.*, 1996), have shown that only Phe and Asn provide weak alternatives to the Tyr75 in stabilizing the intermediate/transition state of the catalytic mechanism.

We have chosen to work on chymosin as this is an enzyme of interest to the food industry and the recombinant enzyme is now used widely in food processing (Flamm, 1991). Bovine chymosin is found in the fourth stomach of the suckling calf (Foltmann, 1981). Its active enzyme is formed by proteolytic cleavage at acidic pH of the N-terminal 42-residue propeptide of its zymogen, prochymosin. Chymosin preferentially cleaves the milk protein κ-casein at the Phe105–Met106 bond:

-His-Pro-His-Pro-His-Leu-Ser-Phe—Met-Ala-Ile-Pro-Pro-Lys-Lys-
98 105 ↑ 106 112
Cleavage Site

inducing instability in milk micelles, which leads to clotting.

The sequence of bovine chymosin (Foltmann *et al.*, 1979) resembles that of other mammalian enzymes such as pepsin, cathepsin D and kidney renin (Kostka, 1985) with two aspartic acid residues at 32 and 215 respectively (pepsin numbering) that are essential for catalytic activity. Calf chymosin is found in three main forms, A, B and C. Chymosin A and B have intact chains but differ in the residue at 243 (pepsin numbering): glycine is present in chymosin B while aspartic acid occurs in chymosin A. Activity measurements indicate that chymosin A has a 20% greater milk clotting ability than chymosin B (Foltmann, 1966). Chymosin C appears to be a degradation product of chymosin A (Danley and Geoghegan, 1988).

Although crystals of chymosin were among the first to be studied by X-ray analysis (Bunn *et al*., 1971), three-dimensional structures of the native mature chymosin were reported much later at 2.3 Å by Gilliland *et al.* (1990) and at 2.2 Å by Newman *et al*. (1991). The active site 'flap' that covers the cleft and interacts with the substrate is rearranged in the native crystals, preventing access of substrate or inhibitor. The purification and crystallization of chymosin G243D expressed from *T.reesei* has been previously reported (Pitts *et al.*, 1991; 1992; 1993). Recently we have obtained a new crystal form of chymosin complexed with a small inhibitor and we have solved the structure at 2.3 Å resolution (Groves *et al*., 1997). This analysis shows that the inhibitor binding is quite normal; it is not possible to define whether the rearranged structure in the native crystals is a consequence of crystal packing or occurs in solution.

We have developed site-directed mutagenesis studies of chymosin to investigate the role of individual amino acids in the determination of the pH optimum (Mantafounis and Pitts, 1990; Pitts and Mantafounis, 1991; Pitts *et al.*, 1992), specificity (Quinn *et al.*, 1991) and stability. We have made a detailed three-dimensional structure and kinetic analysis of one mutant Val111Phe showing that changes in kinetic parameters can be explained in terms of detailed structural changes at specificity subsites (Strop *et al*., 1990).

In this paper we describe the site-directed mutagenesis, kinetic characterization and crystal structure analysis of two point mutants of chymosin B, chymosin G243D (chymosin A) and chymosin A115T, where changes introduced on the periphery of the active site cleft are used to examine the role of these regions in changing specificity and kinetics. These mutants were two of many introduced in the periphery of the active site. To some extent they were self selected from amongst a large group by their ability to crystallize. Most mutants did not crystallize, but where crystals were obtained, they were often stacks of misorientated plates or too small to allow collection of diffraction data. Chymosin G243D was chosen because it is a natural mutant occurring in chymosin B; although it is at the periphery, earlier studies on whole milk had indicated that it affects activity. Chymosin A115T was one of a large number of mutations in a helical region adjacent to subsites S1 and S3 which were considered to have a strong possibility of influencing the activity. To facilitate the comparisons, we use the pepsin numbering system (Barkholt, 1987; Blundell *et al*., 1990). We describe the structure of bovine chymosin G243D at 2.8 Å resolution and of chymosin A115T at 2.3 Å resolution, and we compare their structures and kinetics to those of the native enzyme.

Materials and methods

Organisms and plasmids

Chymosin G243D was expressed in *Trichoderma reesei* strain RutC, using the expression vector pAMH104, as described by Harkki *et al.* (1989). During growth of *Trichoderma* on the inducing carbon source cellulose, folded and active enzyme was secreted into the culture medium.

Chymosin A115T was expressed as inclusion bodies in *E.coli* MT [F′, *pro*AB, LacI^q] derived from *E.coli* HB101 (Kapralek *et al.*, 1991). The expression plasmid, pMG1225, was identical to the plasmid pMG225 already described (Kapralek *et al.*, 1991; Tichy *et al.*, 1993) except for the A115T mutation.

Isolation of chymosin G243D

Spent fermentation medium containing secreted chymosin G243D was clarified and spray dried, prior to refrigerated storage. Spray-dried spent medium (107 g) was resuspended in distilled water (800 ml) and stirred for 1 h at room temperature. The resulting suspension was clarified by centrifugation (8000 r.p.m., 30 min) and filtration (0.6 m glass fibre filters), chilled to 4°C, and then adjusted to 40% saturation with solid ammonium sulphate. After stirring overnight at 4°C, precipitated protein was isolated by centrifugation (5000 r.p.m., 30 min, 5°C), and redissolved in 0.1 M Na acetate buffer, pH 5.6, 0.1 M NaCl (buffer A: 300 ml).

Redissolved protein in buffer A was fractionated in two equal aliquots, by gel filtration chromatography on a Sephadex G100 column (5 cm×90 cm) equilibrated in buffer A. The column was developed at 2 ml/min, and fractions (11 ml) collected. The protein elution profile was monitored at 280 nm, and the relative chymosin G243D content of each fraction determined by a milk-clotting assay (see below).

Fractions containing milk-clotting activity from both G100 runs were pooled, and concentrated by diafiltration using an Amicon stirred cell and PM10 membrane (10 000 kDa cut-off). Buffer A was exchanged for 10 mM Tris, 10 mM acetate buffer (buffer B) by repeated concentration and dilution, and the pooled chymosin G243D applied to a HFQ cation exchange column (5 cm×20 cm) equilibrated in buffer B, and developed at 10 ml/min. Proteins were fractionated by the application of a gradient of 0–0.6 M NaCl in buffer B. The protein elution profile was again monitored at 280 nm, and the relative chymosin G243D content of each fraction determined by milk clotting.

Fractions containing milk-clotting activity were again pooled, concentrated by diafiltration as above to 30 ml, and subjected to gel filtration chromatography on a Sephacryl S100 column equilibrated in buffer A, under the same conditions as above. The protein elution profile was again monitored at 280 nm, and the relative chymosin G243D content of each fraction determined by milk clotting. Fractions containing milk-clotting activity were pooled, and the enzyme solution used for either crystallization experiments or kinetic analysis. The purity of isolated recombinant enzyme was confirmed by SDS–PAGE (results not presented) in comparison with chymosin isolated from rennet (Foltmann, 1970). Typically, approximately 35 mg enzyme was obtained from 100 g dried spent medium.

Cultivation of E.coli MT and isolation of chymosin A115T

Escherichia coli MT cells harbouring pMG1225 were grown at 37°C in culture medium containing tryptone, yeast extract and lactose; aerobic cultivations used Erlenmayer flasks placed on a rotary shaker, or a laboratory fermenter (Magnaferm MA114, New Brunswick).

Formation of inclusions of the prochymosin product in the *E.coli* cells was quantitatively evaluated under phase contrast microscopy by determining the diameter of an average inclu-

sion body and calculating the volume V_i of inclusions in 1 ml of the culture, the volume V_b of the cells present in 1 ml of the culture, and the ratio V_i/V_b (Kapralek *et al.*, 1991).

In our *E.coli* expression system recombinant prochymosin is accumulated as inclusion bodies. The isolation of prochymosin from inclusion bodies uses the procedure of Marston *et al.* (1984) with modifications described previously (Tichy *et al.*, 1993). Wet biomass (20 g) was processed by standard cell lysis procedures comprising lysozyme/sodium deoxycholate treatment and ultrasonic disintegration (Marston *et al.*, 1984).

Inclusions pelleted by centrifugation (6000 *g*, 5 min, 5°C) were washed in buffer A (50 mM Tris–HCl, pH 8.0, 1 mM EDTA, and 50 mM NaCl) containing 0.5% (v/v) Triton X-100 and EDTA increased to 10 mM final concentration. Washed inclusions were collected by centrifugation as before. Their solubilization in 8 M urea solution (in 50 mM KH_2PO_4, pH 10.4, 1 mM EDTA, 50 mM NaCl) and prochymosin renaturation by diluting into an alkaline buffer (pH 10.7) was according to optimized procedure of Tichy *et al.* (1993). The urea/alkali solution was then diluted approximately two times with distilled water to the conductivity of buffer B (20 mM Tris–HCl, pH 8.0, 1 mM EDTA, 50 mM NaCl) and applied to a DEAE-cellulose column (1.8×10 cm) equilibrated in the above buffer. The column was washed with four volumes of 20 mM Tris–HCl, pH 8.0, 1 mM EDTA, 100 mM NaCl, followed by elution of the prochymosin material with a step to 500 mM NaCl in four volumes.

Fractions (4 ml) were collected, 0.05 ml samples were withdrawn and after an activation step (acidification to pH 2.0, as described in Emtage *et al.*, 1983) assayed for chymosin activity by a modified turbidimetric milk-clotting assay of McPhie (1976). Fractions containing activatable prochymosin product formed two close peaks in the elution profile. Fractions of the second peak were separately activated and pooled. The combined activation mixtures (about 100 ml) were concentrated twofold by ultrafiltration using a 10 000 MW cut-off membrane (Amicon, USA, type YM10) and then applied to a Mono Q 5/5 column (Pharmacia, Uppsala) equilibrated with 50 mM KH_2PO_4, pH 6.5, 1 mM EDTA and 0.01% sodium azide. The column was eluted for 40 min at a flow rate of 1 ml/min with a linear gradient to 1.3 M NaCl chymosin eluted as a single active peak giving a final yield of purified enzyme of 76 mg from 20 g of wet biomass.

Kinetic analysis

For kinetic analysis, chymosins were further purified by affinity chromatography using a protocol modified from Strop *et al.* (1990), described previously (Nugent *et al.*, 1996).

Two synthetic peptide substrates were employed. The first substrate Leu-Ser-Phe(NO_2)-Nle-Ala-Leu-OMe was purchased from Bachem (Saffron Walden, UK). The second, His-Pro-His-Pro-His-Leu-Ser-Phe-Phe(NO_2)-Ala-Ile-Pro-Pro-Lys-Lys, based on residues 98–112 of κ-casein, was synthesised on an ABI 431A peptide synthesizer using Fmoc chemistry, and purified by reverse phase chromatography on a C18 column developed with a gradient of acetonitrile in 0.1% TFA.

Kinetic measurements were performed in 0.1 M acetate buffer, pH 5.6, 1 mM $CaC1_2$, 30°C, and the change in absorbance at 305 nm was followed. In the absence of a strongly binding, specific inhibitor to chymosin to allow the active site to be titrated, the k_{cat} values were based on the concentration of enzyme estimated by either Lowry assay, or a E1%1cm 280nm value of 14 (Rothe *et al.*, 1976).

Table I. Structure determination summary

	Chymosin G243D from *T.reesei*	Chymosin A115T from *E.coli*
Unit cell dimensions		
a	79.64 Å	80.47 Å
b	113.73 Å	114.15 Å
c	72.40 Å	72.59 Å
α = β = γ	90°	90°
Space group	I222	I222
Number of molecules in asymmetric unit	1	1
λ	0.92 Å	1.5418 Å
R_{sym}[a]	0.138	0.094
Number of unique reflections	6442	14719
% measured/possible	77.5	98.4
% I > 3σ (I)	71.9	– – – – –
R_{cryst}[b](%)	18.8	17.9
Structure solution	Isomorphous to native crystals	Isomorphous to native crystals
Refinement	RESTRAIN/X-PLOR	RESTRAIN
Resolution (Å)	15–2.8	17–2.3
Non-hydrogen atoms	2515 + 77 H_2O	2513 + 171 H_2O
RMS bonds (Å)	0.008[c]	0.014[d]
RMS angles (°)	1.998[c]	0.028[d]

[a]$R_{sym} = \Sigma_h\Sigma_i|<F_h>-F_{hi}|/\Sigma_hF_h$, where $<F_h>$ is the mean structure factor amplitude of the *i* observations of reflections that are related to the Bragg index h.
[b]$R_{cryst} = \Sigma||F_o|-|F_c||/\Sigma|F_o|$, where F_o and F_c are the observed and calculated structure factor amplitudes respectively.
[c]R.m.s. deviations of bond lengths and angles in the refined structure from ideal values (Engh and Huber, 1991).
[d]R.m.s. deviations from target bond length and angle values in RESTRAIN (Haneef *et al.*, 1985).

Crystallization and X-ray analysis

Crystals suitable for X-ray analysis of both sets of proteins were grown by the hanging-drop vapour diffusion method (McPherson, 1982). Chymosin G243D and A115T grew using the same crystallization protocols: 10 µl protein solution at a concentration of 8 mg/ml in 1.0 M NaCl solution and 50 mM phosphate buffer at pH 5.6 were equilibrated against well solutions containing 1.5–1.8 M NaCl in 150–200 mM phosphate buffer of the same pH. Diffraction data were collected for G243D on a MARResearch image plate detector at the Synchrotron Radiation Source, Daresbury, UK, and for A115T on a MARResearch image plate detector, with Ni-filtered Cu-K_α radiation from a rotating anode generator.

The reference orientations of both crystals were established with the programs REFIX and IMSTILLS (Kabsch, 1988), and the program MOSFLM (version 5.1.2) was used to reduce the image plate data, and to refine the data collection and cell parameters. Further merging, scaling, correction and processing of the data were carried out using the CCP4 suite of programs (CCP4, 1994). A summary of the X-ray diffraction data is presented in Table I. Electron density maps for both sets of data could be computed using the chymosin B (PDB code 4CMS) as the starting model with the atomic co-ordinates left unmodified in both cases.

The least-squares refinement of chymosin G243D using the program RESTRAIN (Haneef *et al.*, 1985), with the model treated as a single rigid body, and subsequently with three separate rigid body domains (residues –2 to 185; 186 to 299; and 300 to 326) led to an *R* factor of 0.30. Subsequent all-atom positional refinement, with an overall temperature factor

Table II. Kinetic analysis of recombinant chymosin B (rCMSB), chymosin A115T (rA115T) and chymosin G243D (rG243D)

Enzyme	K_m (μM)	V_{max} (μM min^{-1})	k_{cat} (s^{-1})	k_{cat}/K_m ($\mu M^{-1}s^{-1}$)
Kinetic parameters using Leu-Ser-Phe(NO_2)-Nle-Ala-Leu-OMe substrate				
rCMSB	277 ± 4.9	10.3 ± 0.25	12.9 ± 0.31	0.047 ± 0.001
rA115T	1110 ± 19.6	10.3 ± 0.25	14.1 ± 0.34	0.013 ± 0.002
rG243D	333 ± 1.0	5.7 ± 0.035	4.8 ± 0.013	0.014 ± 0.001
Kinetic parameters using κ-casein substrate at pH 5.3				
rCMSB	11.0 ± 0.104	4.5 ± 0	94.2 ± 0	8.6 ± 0.12
rG243D	20.0 ± 0.07	10.0 ± 0.477	208.5 ± 3.69	10.4 ± 0.015

Parameters determined using Lineweaver–Burke plot (1/*v* versus 1/[*S*]).

for the model, using a data shell from 8.0 to 2.8 Å resolution, gave an improved *R* value of 0.24. F_o–F_c and $2F_o$–F_c electron density maps, contoured at ±2 r.m.s. and +0.9 r.m.s. respectively, allowed remodelling using FRODO (Jones, 1978) on an Evans and Sutherland PS390 graphics terminal, followed by refinement and rebuilding. At this stage the N-terminus and the loops between residues 9–13, 154–164, 240–244 and 278–281 were discarded. Initially the density for residue 243 did not show clear evidence of a sidechain for the aspartic acid residue, and there was no density for the three residues 290–292, which had been omitted from the co-ordinates of 4CMS deposited at the Brookhaven data bank. Re-modelling residues 290–292 following the conformation found in the chymosin–inhibitor complex of Groves *et al.* (1997) and re-allocation of water molecules, followed by simulated annealing in the program X-PLOR (Brunger, 1990) using force-field parameters derived by Engh and Huber (1991) for data in the 15.0 to 2.8 Å resolution range resulted in the value of *R* dropping to 0.22, and after positional refinement to 0.21. In the subsequent calculated $2F_o$–F_c and F_o–F_c, now contoured at +1 and ±2 r.m.s. respectively, all the regions showed improvement including the previous disordered regions. Density was observed around the Asp243 residue and almost completely along the C_α chain of the loop 240–244 with density appearing in the region of residues 290–292. These improvements prompted the re-modelling of four regions with the weakest electron density. Water molecules were calculated by automatic peak-picking within the complete F_o–F_c difference map at ±2 r.m.s. using the programs PEAKMAX and WATPEAK (CCP4, 1994). Only molecules that gave sensible hydrogen bonds to the protein were incorporated into the model and their geometries checked with the program CONTACT (CCP4, 1994). In subsequent rounds of refinement, waters with high temperature factors ($B > 45$) were removed. The last simulated annealing refinement with the modified loop conformations refined to give a value for *R* of 0.20, further positional and grouped *B*-factor refinement gave a final *R* of 0.19.

The chymosin A115T was also refined using the program RESTRAIN using co-ordinates of the native chymosin B structure (4CMS) with Ala115 and the water structure both unmodified. Rigid body refinement, followed by all-atom positional refinement, and ultimately including refinement of individual isotropic temperature factors with all data to 2.3 Å led to an *R* value of 18.5% without significant manual intervention. At this stage the original water structure was discarded and recalculated by automatic peak-picking within the difference map using the same procedure as above, with the exception that water molecules with high temperature factors ($B > 75$) were deleted.

In the original native co-ordinates (4CMS), three residues were omitted completely (290–292), and three loops (236–244, 274–279 and 288–292) were poorly defined probably due to loop mobility. This A115T mutant structure had similar problems, with poor map definition and high *B*-factors in the region of these loops, as well as in the region of residues 10–16. The final *R* value for all data from 17.0 to 2.31 Å was 0.178. In the final model 168 water molecules are included, some of which are buried within the protein, as was the case in the native structure.

Substrate modelling in chymosin G243D

In order to assess the likely effect of the mutation G243D on the activity of chymosin towards, κ-casein, a model of the chymosin substrate transition state complex was made using a fifteen residue fragment. The conformation was based on the structure of a complex of the inhibitor with chymosin B (Groves *et al.*, 1997), on the structures of inhibitors complexed with endothiapepsin (Bailey and Cooper, 1994), and on the structure of a long peptide (P6-P4′) complexed with mouse renins (Dhanaraj *et al.*, 1992; DeAlwis *et al.*, 1994). There were few structural restraints on the remainder of the residues of the substrate. Residues 74–79 of the 'flap' region were replaced with the conformation of the corresponding region in the chymosin inhibitor complex (Groves *et al.*, 1997).

The substrate fragment was positioned within the active site of chymosin G243D, with the help of the structure of the inhibitor complex as a guide, using the program O (Jones *et al.*, 1991). SYBYL (version 6.1; Tripos Associates Inc.) was used to minimize bad contacts in the model which had occurred between atoms in the cleft of the chymosin G243D and κ-casein side-chains. Residues (98–102) of κ-casein were modelled interacting with the 240–244 loop.

Results and discussion

Kinetic measurements

In our kinetic analysis we have used two peptides. The short peptide has been used by others to characterize the enzyme activity of chymosin. Other workers can, therefore, easily compare our results with those reported by other investigators. The use of the longer κ-casein peptide substrate allowed us to probe for changes in enzyme structure, upon mutation, which affect interactions with the substrate at residues more distal to the scissile bond of the substrate.

Comparison of the kinetic characteristics (see Table II) of chymosin A115T, chymosin G243D and wild-type chymosin B, using the hexapeptide substrate, showed that the mutations caused contrasting alterations in enzymatic activity (see Figure 1a). The A115T mutation reduced the substrate binding ability (K_m) of the enzyme but not its catalytic efficiency (k_{cat}), while the G243D mutation reduced the catalytic efficiency of

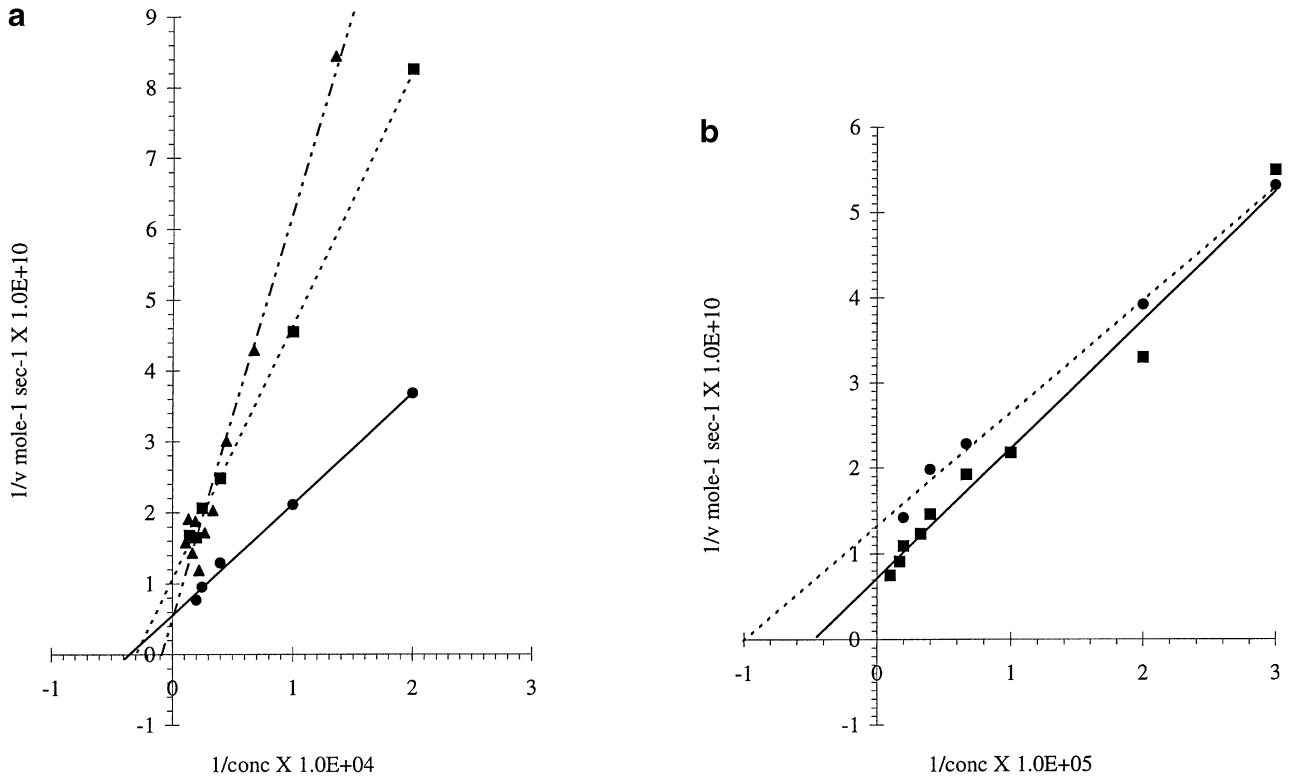

Fig. 1. Lineweaver–Burke plots of (**a**) recombinant chymosin B (●; r = 0.996), chymosin G243D (■; r = 0.999) and chymosin A115T (▲; r = 0.969), with a hexapeptide substrate; and (b) recombinant chymosin B (●; r = 0.982) and chymosin G243D (■; r = 0.995) with a 15 residue κ-casein peptide. The straight lines through the points represent a linear regression, with the goodness of fit represented in parentheses.

the enzyme but had little effect on its substrate binding ability. In both cases, mutation significantly reduced the specificity constant (k_{cat}/K_m) of the enzyme.

Comparison of wild type chymosin B with chymosin G243D (chymosin A), using κ-casein as substrate (see Figure 1b) reflected earlier observations (Rothe *et al.*, 1976) that chymosin G243D has higher activity towards milk. The G243D mutation resulted in a small reduction in the substrate binding ability of the enzyme, but increased its catalytic efficiency. However, the specificity constant of chymosin G243D was not significantly different to that of the wild-type chymosin B enzyme.

Comparison of the kinetic parameters determined with these two substrates clearly demonstrates the superiority of the 15 residue κ-casein peptide over the shorter six residue peptide as a substrate for chymosin.

X-ray structures

Both materials crystallized in the orthorhombic space group, I222, with cell dimensions a = 79.64, b = 113.73, c = 72.40 Å, α = β = γ =90° for chymosin G243D, and a = 80.47, b = 114.15, c = 72.59 Å, α = β = γ =90° for chymosin A115T, isomorphous to the crystals of the native 'wild-type' chymosin B (Newman *et al.*, 1991), with a single molecule per asymmetric unit.

The density at residue 115 in both $2F_o–F_c$ and $F_o–F_c$ maps, contoured at +1 and ±3 r.m.s. respectively, was consistent with the mutation of alanine at 115 to a threonine. This region gave a peak of positive difference density at the position expected for the side-chain of a threonine. Negative difference density indicated that there had been a slight displacement of the helix between residues 112 and 117, probably due to the extra methyl group and γ-oxygen in the mutant structure. Building the threonine side-chain into the maps permitted the introduction of a new hydrogen bond between its γ-oxygen and Phe112 main-chain carbonyl. The water molecule (OW347) from the chymosin B structure was then seen to have been sterically occluded by the addition of the side-chain atoms.

The structure of recombinant chymosin G243D is very similar to that of chymosin B (4CMS) and (1CMS). The loops 8–13, 158–163, 240–244 and 289–297 have weak density although it is continuous at +1 RMS between residues 239–243 and between 289–291 and 292–297. These poorly defined loop regions on the surface of the protein are also evident in the *B*-values shown.

The main differences in the C_α positions of chymosin G243D and the two 'wild-type' chymosin B structures occur in the loop 278–281 and unsurprisingly, in loop 240–244, where the only sequential difference in the protein arises. The carboxylate oxygen of Asp243 forms a hydrogen bond with the NE2 nitrogen on the Gln239 side-chain (see Figure 2a and b), and with a water molecule (HOH 1050). The loop (240–244) has a significantly different conformation.

It has been postulated that the increased activity of chymosin G243D may be the result of an enhanced binding affinity for its natural substrate, κ-casein, due to the resultant electrostatic interactions between His102 (P4) of the substrate and Asp243 of the enzyme (Gilliland *et al.*, 1991). However, from the results of modelling the κ-casein in the chymosin G243D structure by extending the peptide (P8–P7′), P4 His is seen to be approximately 12 Å away from Asp243. Since positions P5 to P8 in the modelled substrate are not rigidly bound in the enzyme, it is possible to model the κ-casein sequence, His-Pro-His-Pro (residues 98–101, P8–P5), so that it interacts with the mutated residue (see Figure 3). His98 (P8) is within 3.2 Å of OD2 in Asp243, while Pro99 (P7) and His100 (P6)

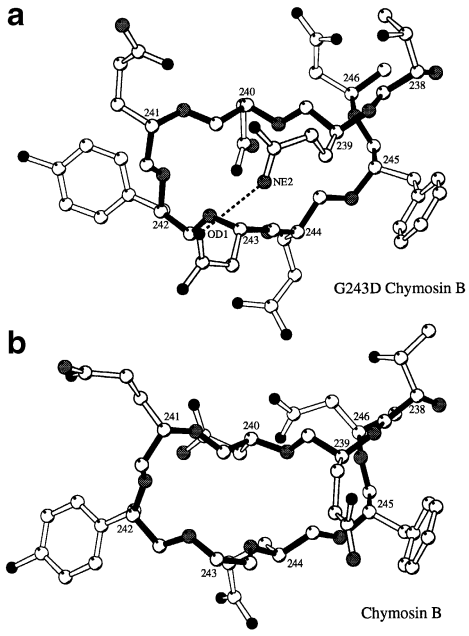

Fig. 2. Representation of residues 238–246 of G243D chymosin B, showing (**a**) the hydrogen bond between NE2 of Gln239 and OD1 of Asp243; and (**b**) viewed in an identical orientation as above, for native chymosin B (4CMS). Both figures were generated using the program MOLSCRIPT (Kraulis *et al.*, 1991).

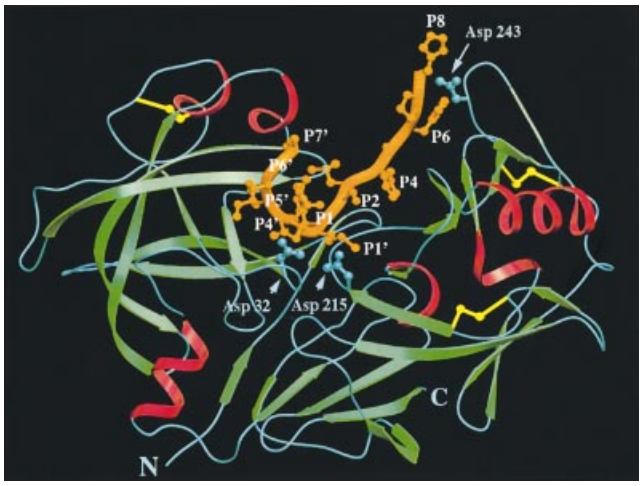

Fig. 3. A pictorial representation of the model of the κ-casein substrate transition state complex with G243D chymosin B generated using the program SETOR (Evans, 1993).

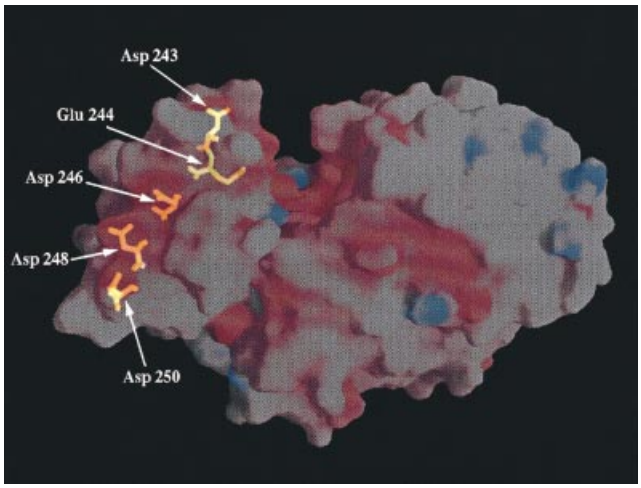

Fig. 4. A representation generated using the program GRASP (Nicholls *et al.*, 1993) of the relative charge distribution (red, negative; blue, positive) on the surface of the protein. The positions of residues Asp243, Glu244, Asp246, Asp248 and Asp250 are indicated.

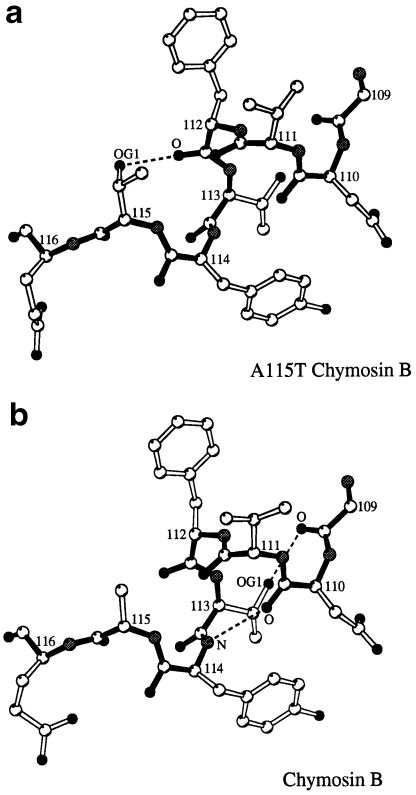

Fig. 5. Residues 109–116 of (**a**) the A115T chymosin B mutant, showing the hydrogen bond between O of Phe112 and OG1 of Thr115; and (**b**) in an identical orientation to the above, native chymosin B (4CMS), with hydrogen bonds between O of Asp110 and N of Tyr114, and between O of Gly109 and OG1 of Thr113. Both figures generated using MOLSCRIPT.

come close to the loop (240–244), and His102 (P4) is about 5.5 Å away from OE2 in Glu244. His100 is also 3.5 Å from Asp278 which could mean that the κ-casein sequence (residues 98–102), His-Pro-His-Pro-His, provides additional binding energy, with the histidines interacting electrostatically with the loop (278–281) of chymosin G243D.

Asp243 is positioned in a loop on the surface of the enzyme and is very close to additional negatively charged residues, Glu244, Asp246, Asp248 and Asp250 (see Figure 4). This additional negative charge should aid the stabilization of the substrate–enzyme complex. The increase in enzymatic activity for chymosin G243D over chymosin B towards κ-casein (as found in whole milk) could, therefore, be the result of favourable interactions with large peptides.

The A115T mutant chymosin produced in *E.coli*, altered only by the addition of a methyl group and an oxygen to a single side-chain, is largely unchanged compared with the native structure (see Figure 5a and b). The Thr115 side-chain is accommodated with small local movements, and forms a new side-chain to main-chain (115OG1 → 112O) hydrogen bond. Residue 115 is at the end of a single turn of irregular helix, and close comparison of this region in the mutant and native structures reveals that the steric accommodation of the main-chain and side-chain in the region of Phe112 in the

mutant structure involves a shift of the order of 0.6 Å. This shift extends over 10 or more residues back along the chain, with the loss of two hydrogen bonds at the single turn of the helix (both 113OG1 → 109O and 114N → 110O are lost). A slight tilting of the 'flap' (residues 72–82) by about 0.5 Å is also discernible, consequent on the 'upwards' movement of the bulky phenylalanyl-112 side-chain. These small shifts are likely to affect the specificity pockets at S1 and S3 which involve residues on this helix.

Acknowledgements

We thank the BBSRC for financial support of this program and Dr D. Burke for helpful discussions on the substrate modelling.

References

Bailey,D. and Cooper,J.B. (1994) *Protein Sci.*, **3**, 2129–2143.
Barkholt,V. (1987) *Eur. J. Biochem.*, **167**, 327–338.
Blundell,T.L., Cooper,J.B., Tickle,I.J., Veerapandian,B. and Wood,S.P. (1990) *J. Mol. Biol.*, **210**, 919–941.
Bott,R., Subramanian,E. and Davies,D.R. (1982) *Biochemistry*, **21**, 6956–6962.
Brodelius,P.E., Cordeiro,M.C. and Pais,M.S. (1993) *Proc. of 5th International Conference on Aspartic Proteinases*, L14 and P30.
Brunger,A.T. (1990) *X-PLOR version 3.1. A System for X-ray Crystallography and NMR*. Yale University Press, New Haven, CT.
Bunn,C.W., Moews,P.C. and Baumber,M.E. (1971) *Proc. Roy. Soc. Lond.*, **B178**, 245–258.
CCP4 (1994) *Acta Crystallogr.*, **D50**, 760–763.
Cordeiro,M.C., Xue,Z.T., Pietrzak,M., Pais,M.S. and Brodelius,P.E. (1994) *Plant Mol. Biol.*, **24**, 733–741.
Danley,D.E. and Geoghegan,K.F. (1988) *J. Biol. Chem.*, **263**, 9785–9789.
Davies,D.R. (1990) *Annu. Rev. Biophys. Chem.*, **19**, 189–215.
DeAlwis,C.G., Frazao,C., Badasso,M., Cooper,J.B., Tickle,I.J., Driessen,H., Blundell,T.L., Murakami,K., Miyazaki,H., Sueiras-Diaz,J., Jones,D.M. and Szelke,M. (1994) *J. Mol. Biol.*, **236**, 342–360.
Dhanaraj,V., DeAlwis,C., Frazao,C., Badasso,M., Sibanda,B.L., Tickle,I.J., Cooper,J.B., Driessen,H.P.C., Newman,M., Aguilar,C., Wood,S.P., Blundell,T.L., Hobart,P.M., Geohegan,K.F., Ammirati,M.J., Danley,D.E., O'Connor,B.A. and Hoover,D.J. (1992) *Nature*, **357**, 466–472.
Emtage,J.S., Angal,S., Doel,M.T., Harris,T.J.R., Jenkins,B., Lilley,G. and Lowe,P.A. (1983) *Proc. Natl Acad. Sci. USA*, **80**, 3671–3675.
Engh,R.A. and Huber,R. (1991) *Acta Crystallogr.*, **A47**, 393–400.
Evans,S.V. (1993) *J. Mol. Graphics*, **11**, 134–138.
Flamm,E.L.(1991) *Bio-Technology*, **9**, 349–351.
Foltmann,B. (1966) *Compt. Rend. Trav. Lab. Carlsberg*, **35**, 143–231.
Foltmann,B. (1970) *Methods Enzymol.*, **19**, 421–436.
Foltmann,B., Pederson,V.B., Kauffman,D. and Wybrandt,G. (1979) *J. Biol. Chem.*, **254**, 8447–8456.
Foltmann,B. (1981) *Essays Biochem.*, **17**, 52–84.
Foundling,S.I., Cooper,J.B., Watson,F.E., Pearl,L.H., Sibanda,B.L., Wood,S.P., Blundell,T.L., Valler,M.J., Norey,C.G., Kay,J., Boger,J., Dunn,B.M., Leckie,B.J., Jones,D.M., Atrash,B., Hallett,A. and Szelke,M. (1987) *Nature*, **327**, 349–352.
Gilliland,G.L., Winborne,E.L., Nachman,J. and Wlodawer,A. (1990) *Proteins*, **8**, 82–101.
Gilliland,G.L., Toner Oliva,M. and Dill,J. (1991) In Dunn,B.M. (ed.), *Structure and Function of the Aspartic Proteinases*. Plenum Press, New York, 23–37.
Gustchina,E., Rumsh,L., Ginodman,L., Majer,P. and Andreeva,N. (1996) *FEBS Lett.*, **379**, 60–62.
Groves,M.R., Dhanaraj,V., Badasso,M., Pitts,J.E., Hoover,D.J. and Blundell,T.L. (1997) *Prot. Engng.*, in press.
Haneef,I., Moss,D.S., Stanford,M.J. and Borkakoti,N. (1985) *Acta Crystallogr.*, **A41**, 426–433.
Harkki,A., Uusitalo,J., Bailey,M., Penttila,M. and Knowles,J.K.C. (1989) *Bio-Technology*, **7**, 596–603.
James,M.N.G., Sielecki,A.R., Salituro,F., Rich,D.H. and Hofmann,T. (1982) *Proc. Natl Acad. Sci. USA*, **79**, 6137–6142.
James,M.N.G., Sielecki,A.R., Hayakawa,K. and Gelb,M.H. (1992) *Biochemistry*, **31**, 3872–3886.
Jones,T.A. (1978) *J. Appl. Crystallogr.*, **11**, 268–272.
Jones,T.A., Zou,J.Y., Cowan,S.W. and Kjeldgaard,M. (1991) *Acta Crystallogr.*, **A47**, 110–119.
Kabsch,W. (1988) *J. Appl. Cryst.*, **21**, 67–71.
Kapralek,F., Jecmen,P., Sedlacek,J., Fabry,M. and Zadrazil,S. (1991) *Biotechnol. Bioeng.*, **37**, 71–79.
Kostka,V. (1985) *Aspartic Proteinases and their Inhibitors*, Walter de Gruyter, Berlin.
Kraulis,P.J. (1991) *J. Appl. Crystallogr.*, **24**, 946–950.
Lowther,W.T., Majer,P. and Dunn,B.M. (1995) *Protein Sci.*, **4**, 689–702.
Mantafounis,D. and Pitts,J. (1990) *Protein Engng.*, **3**, 605–609.
Marston,F.A.O., Lowe,P.A., Doel,M.T., Schoemaker,J.M., White,S. and Angal,S. (1984) *Bio-Technology*, **2**, 800–804.
McPhie,P. (1976) *Anal. Biochem.*, **73**, 258–261.
McPherson,A. (1982) *Preparation and Analysis of Protein Crystals*. John Wiley and Sons, New York.
Newman,N., Safro,M., Frazao,C., Khan,G., Zdarov,A., Tickle,I.J., Blundell,T.L. and Andreeva,N.S. (1991) *J. Mol. Biol.*, **221**, 1295–1309.
Newman,M., Watson,F., Roychowdhury,P., Jones,H., Badasso,M., Cleasby,A., Wood,S.P., Tickle,I.J. and Blundell,T.L. (1993) *J. Mol. Biol.*, **230**, 260–283.
Nicholls,A., Bharadwaj,R. and Honig,B., (1993) *Biophys. J.*, **64**, A166–A166.
Nugent,P.G., Albert,A., Orprayoon,P., Wilsher,J., Pitts,J.E., Blundell,T.L. and Dhanaraj,V. (1996) *Protein Engng.*, **9**, 885–893.
Park,Y.N., Aikawa,J., Nishiyama,M., Horinouchi,S. and Beppu,T. (1996) *Prot. Engng.*, **9**, 869–875.
Pitts,J.E. and Mantafounis,D. (1991) *Biochem. Soc. Trans.*, **19**, 266.
Pitts,J.E., Quinn,D., Uusitalo,J.M. and Penttila,M., (1991) *Biochem. Soc. Trans.*, **19**, 663–666.
Pitts,J.E., Dhanaraj,V., Dealwis,C.G., Mantafounis,D., Nugent,P.G., Orprayoon,P., Cooper,J.B., Newman,M. and Blundell,T.L., (1992) *Scand. J. Clin. Lab. Invest.*, **52**, 39–50.
Pitts,J.E., Uusitalo,J.M., Mantafounis,D., Nugent,P.G., Quinn,D., Orprayoon,P. and Penttila,M. (1993) *J. Biotechnology*, **28**, 69–83.
Quinn,D., Pitts,J.E., Mantafounis,D., Lawler,S.E., Uusitalo,J.M., Penttila,M.E. and Strop,P. (1991) *Biochem. Soc. Trans.*, **19**, 267.
Rao-Naik,C., Guruprasad,K., Batley,B., Rapundalo,S., Hill,J., Blundell,T.L., Kay,J. and Dunn,B.M. (1995) *Proteins*, **22**, 168–181.
Rothe,G.A.L., Axelsen,N.H., Johnk,P. and Foltmann,B. (1976) *J. Dairy Res.*, **43**, 85–95.
Scarborough,P.E. and Dunn,B.M. (1994) *Protein Engng.*, **7**, 495–502.
Scarborough,P.E., Guruprasad,K., Topham,C., Richo,G.R., Conner,G.E., Blundell,T.L. and Dunn,B.M. (1993) *Protein Sci.*, **2**, 264–276.
Schechter,I. and Berger,A. (1967) *Biochem. Biophys. Res. Commun.*, **27**, 157–162.
Strop,P., Sedlacek,J., Stys,J., Kaderabkova,Z., Blaha,I., Pavlickova,L., Pohl,J., Fabry,M., Newman,M., Frazao,C., Shearer,A., Tickle,I.J. and Blundell,T.L. (1990) *Biochemistry*, **29**, 9863–9871.
Suguna,K., Padlan,E.A., Bott,R., Boger,J., Parris,K.D. and Davies,D. (1992) *Proteins*, **13**, 195–205.
Suzuki,J., Sasaki,K., Sasao,Y., Hamu,A., Kawasaki,H., Nishiyama,M., Horinouchi,S. and Beppu,T. (1989) *Protein Engng.*, **2**, 563–569.
Tichy,P.J., Kapralek,F. and Jecmen,P. (1993) *Protein Exp. Purific.*, **4**, 59–63.
Voigt,J., Kamaruddin,S., Heinrichs,H., Wrann,D., Senyuk,V. and Biehl,B. (1995) *J. Plant Physiol.*, **145**, 299–307.

Received January 2, 1997; revised May 21, 1997; accepted May 27, 1997